DOCUMENT RESUME

ED 298 074 SO 019 575

AUTHOR	Schenk, Robert E.
TITLE	Economics: An Analysis of Unintended Consequences. Volume 1: Introduction to Microeconomics.
PUB DATE	88
NOTE	252p.; For volume 2, see SO 019 576.
PUB TYPE	Guides - Classroom Use - Materials (For Learner) (051) -- Guides - Classroom Use - Guides (For Teachers) (052)
EDRS PRICE	MF01/PC11 Plus Postage.
DESCRIPTORS	Course Content; Curriculum Guides; *Economics; *Economics Education; High Schools; Instructional Materials; Monetary Systems; Social Studies
IDENTIFIERS	*Economic Theory; *Microeconomics

ABSTRACT

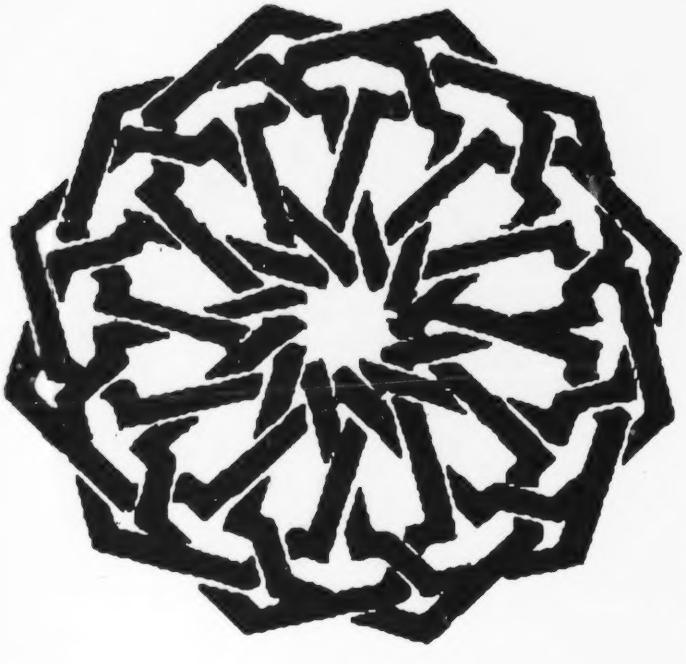
This curriculum guide introduces high school students to the basic principles of microeconomics. Chapter 1 provides a basic definition of economics, while chapter 2 introduces a number of important economic concepts and ideas and examines reasons for unintended or unexpected consequences of decision-making. Chapter 3 considers how individual motives and group results are related, and a model of supply and demand is described in chapter 4. Chapters 5-16 present: (1) elasticity and revue; (2) rationing and allocating; (3) the logic of choice; (4) maximizing behavior; (5) the business firm and its constraints; (6) maximizing profit; (7) efficiency; (8) problems of information, risk, and exclusion; (9) the government and economic efficiency; (10) problems in resource markets; (11) monopoly and efficiency; and (12) transport costs. Tables and charts are included. (JHP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED298074

ECONOMICS

AN ANALYSIS OF UNINTENDED CONSEQUENCES



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality
● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

VOLUME 1

INTRODUCTION TO MICROECONOMICS

SO 019575

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
ROBERT E. SCHENK
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ROBERT E. SCHENK
1988

BEST COPY AVAILABLE

ECONOMICS:
AN ANALYSIS OF UNINTENDED CONSEQUENCES

VOLUME ONE

ROBERT E. SCHENK

ECONOMICS: AN ANALYSIS OF UNINTENDED CONSEQUENCES

COPYRIGHT © 1988, 1987, 1986, 1985, 1984, 1983 by Robert E. Schenk. No part of this publication may be reproduced or distributed in any form or by any, means, or stored in a data base or retrieval system, without the prior written permission of the author.

CONTENTS

CHAPTER 1 WHAT IS ECONOMICS? 1

Defining Economics Unintended Consequences and Economics Economics as a Science Positive and Normative Statements How to Approach the Study of Economics Appendix: A Note on Mathematics and Graphs

CHAPTER 2 ACTIONS AND RESULTS 13

Musical Chairs Contingent Behavior Constrained Behavior By-products or Side Effects

CHAPTER 3 THE INDIVIDUAL AND THE GROUP 27

The Prisoners' Dilemma: Two Stories The Assumption of Self-Interest Other Assumptions Two Ways of Coordination Exchange and Politics

CHAPTER 4 THE MODEL OF SUPPLY AND DEMAND 39

Demand Supply The Model of Supply and Demand The Adjustment Process

CHAPTER 5 ELASTICITY AND REVENUE 55

Price Elasticity Computing Price Elasticity Other Elasticities Revenue and the Demand Curve From Elasticity to Marginal Revenue

CHAPTER 6 RATIONING AND ALLOCATING 67

Price as a Rationer Rationing and the Distribution of Income Prices as Incentives and a Communication Network Price and Feedback

CHAPTER 7 THE LOGIC OF CHOICE 83

Goals and Benefits Constraints and Costs The Maximization Principle The Equimarginal Principle The Rationally Ignorant Voter Slavery

CHAPTER 8 MAXIMIZING BEHAVIOR 99

Indifference Curves and Demand In-kind and Cash Transfers Present Value Consumer's Surplus Producer's Surplus Appendix: Indifference Curves and Maximization

CHAPTER 9 THE FIRM AND ITS CONSTRAINTS 115

Circular Flow Dealing with Three Fundamental Tasks The Production Function The Demand Curve for Output Supply of Resources

CHAPTER 10 MAXIMIZING PROFIT 131

Profit The Profit Maximizing Output The Profit-Maximizing Level of an Input Profit Maximizing in Real Firms An Application of the Theory of Profit Maximizing

CHAPTER 11 EFFICIENCY **147**

Definitions Efficiency and Markets Production Efficiency Exchange Efficiency Efficiency between Substitution and Transformation Qualifications

CHAPTER 12 PROBLEMS OF INFORMATION, RISK, AND EXCLUSION **165**

Why Are There Business Organizations? Screening and Signalling Dealing with Risk and Uncertainty Quality and Price Free Riders

CHAPTER 13 THE GOVERNMENT AND ECONOMIC EFFICIENCY **179**

Majority Rule and Economic Efficiency Public Interest and Private Interest The Politics of Taxes The Efficiency of an Excise Tax Rent Seeking

CHAPTER 14 PROBLEMS IN RESOURCE MARKETS **195**

Marginal Productivity and Income What Is Fair? Government Redistribution "Free" Resources

CHAPTER 15 MONOPOLY AND EFFICIENCY **211**

What Is Wrong with Monopoly? The Theory of Few Sellers Price Discrimination Antitrust Policy Regulation

CHAPTER 16 TRANSPORT COSTS AND BORDERS **229**

Hotelling's Model Monopolistic Competition International Trade Barriers to Trade

NOTE TO THE STUDENT

Most students wonder why they should learn anything about the subject matter of the courses they take. They ask, "Will this course make me a better person, or will it help me in a career?" Economics encourages such questioning, for it proposes that people do (and often should) base decisions on the weighing of costs and benefits.

I have not seen any convincing arguments that economics will make you more moral or happier. I doubt if it improves your personality or your social skills. Some economists have suggested that it might make you a better citizen. After taking a course in economics, you should better understand the economic and political processes which shape the world you live in. But even if you have this understanding, there is no guarantee that you will use it for the public good. Knowledge is not goodness. However, economics does tend to form realists rather than idealists, and I suppose that one could argue that a world of realists is a better world than one of idealists.

Neither does economics teach you specific vocational skills which you will be able to apply is some future career. Economics is not like accounting or nursing, which do teach specific skills. Alas, if you want to work as a professional economists, graduate school is necessary.

Though people can live happy and fulfilled lives without economics, there are some things of value which a course in economics does offer the student. After you leave college, many of you will have careers in business or government. You will work in an environment shaped a great deal by economic forces and economic theories. If you master introductory economics, you will develop understandings and methods of thought which will help you cope with that environment.

This book emphasizes that economics is a method of thought or analysis. Many people have found the economic way of thinking very useful. It is used by a great many people who do not consider themselves economists. Lawyers, bankers, bureaucrats, corporate executives, in fact anyone who is in a position in which he or she must make a lot of decisions, can find the economic way of the economic way of thinking useful.

Another way of saying that economics is a method of analysis is to say that it is a technique of problem solving. It cannot solve all problems, but there is a large class of problems for which it is useful. The only way to learn to use a technique of problem solving is to practice it. Hence at the end of each chapter you will find a number of problems which attempt to make you apply the principles of the chapter. Some of them are fairly easy, some are quite difficult. I urge you to spend time and effort on all of them.

Also, notice that each chapter begins with a list of learning objectives. Look at them before you read the chapter, and again after you finish the chapter. You might also look at them in preparing for tests; they make a good study sheet.

TO THE INSTRUCTOR

There are scores of economics textbooks. The world would not seem to need another. Yet the continuing outpouring of new textbooks may indicate a widespread dissatisfaction with existing textbooks. It may also indicate that textbooks, at least introductory textbooks, are fun to write. This book would not have been written had I not been dissatisfied with the textbooks I have seen, and had I not enjoyed the challenge of trying to devise my own.

This book has a number of features which are uncommon in principles microeconomics textbooks. One of the most important is the emphasis given to the problems and questions at the end of each chapter. Some of them are meant to give students practice applying an idea from the chapter and others are included as a way of going a step beyond what is in the chapter. In some cases the student cannot understand important topics unless he or she works some of the problems. This use of problems is intentional. I believe that there is a high return on methods of learning in which students play active roles, and problem solving is one such activity. Learning economics is not a matter of memorization, but of discovering how to approach and solve problems.

More than the problems, however, most instructors will notice the very different organizational structure of this book. There are a lot of economics textbooks, but they mostly have a very standardized structure. Introductory microeconomics stresses market structure. This book de-emphasizes it in order to stress other topics more. Many instructors may feel that departure from this tradition needs justification.

The central theme of this book and, I believe, of economics is the question of how groups of individuals, each acting in his or her own self interest, function. Is the good of the group and the good of the individual in conflict or in harmony? Adam Smith argued that harmony would reign in a market, but only in the absence of monopoly. This way of viewing the market leads to the traditional textbook treatment.

However, the seriousness of the problems which monopoly and its near relatives cause remains a matter of dispute. Some economists think they represent a serious malfunctioning of market economies, while others think their importance less serious. I side with the latter group. I think there are other, more serious problems that market economies have. Given this starting point, the traditional organization, which de-emphasizes problems of information, risk, and property rights by emphasizing market structure, is unacceptable.

A second major problem I find with the traditional text is the way they treat government. Many texts have extensive discussions of government policies, but do not have an explicit theory of government. Such a theory is logically necessary if one wants to know what government can or cannot do. Economics has developed a theory of government, one based on self-interested behavior. Some texts are beginning to add chapters on public choice, but they have tacked it onto the end of the book as an optional chapter, not as an integral part of discussion of government policy. This approach makes no sense. If one wants to discuss government policy, one should begin with a theory of that policy.

Principles textbooks tend to be a conservative force in most disciplines, that is, they tend to preserve ideas even after they have outlived their usefulness. If you are using this book, you obviously are willing to try new approaches. I hope you will enjoy the one I offer.

September 1988

CHAPTER 1

WHAT IS ECONOMICS?

After you read this chapter and work the problems at the end, you should be able to:

Explain what the term "invisible hand" means and who first used it.
Give at least one common definition of economics.
State what Malthus thought about population growth.
Explain how Popper defines scientific statements.
Distinguish between positive and normative statements.
Explain what scarcity, choice, and self-interest have to do with economics.

Economics has been recognized as a special area of study for over a century. Virtually all four-year colleges have course in economics and most allow students to major in the subject. Economists maintain high profiles in governments, and they have been well represented among the highest appointees in the federal government of the United States. The press reports on their doings and sayings, sometimes with praise and admiration, sometimes with ridicule and scorn. Economics and economists are words that almost everyone has heard of and uses. But what exactly is economics? Very few people can give a good definition or description of what this field of study is all about.

DEFINING ECONOMICS

If the ordinary citizen cannot give a good definition or description of economics, he can be excused because economists have also had a hard time defining their field. This difficulty is a problem in all fields which have become complex with fuzzy boundaries and a variety of subdivisions that seem little related. In recent years, for example, economists have written papers and had them published by scholarly economic journals on topics such as sex, crime, slavery, childbearing,, and rats. It is not surprising then that one economist, in a lighter moment, suggested defining economics as "what economists do."

Defining economics as "what economists do" does not tell us anything we did not already know. A good definition must explain what it is that makes economics a distinct subject, different from physics or psychology. One should not expect to find a short definition that conveys with absolute clarity all there is to know about economics (or else there would be no reason to spend hours learning about it). Neither should one believe that there is only one correct definition possible. Many good definitions are possible, and each will focus on some important aspect of the subject. To use an analogy, there is not one spot from which one can best view Niagara Falls. Each viewpoint obscures some features and emphasizes others. There are of course some spots which are clearly superior to others, but people can disagree about which is the very best spot.

One of the earliest and most famous definitions of economics was that of Thomas Carlyle who in the early 19th century termed it the "dismal science." What Carlyle had noticed was the anti-utopian implications of economics. Many utopians, people who believe that a society of abundance without conflict is possible, believe that good results come from good motives and good motives lead to good results. Economists have always disputed this, and it was the forceful statement of this disagreement by early economists such as Thomas Malthus and David Ricardo which Carlyle reacted to.

Another early definition, one which is perhaps more useful, is that of English economist W. Stanley Jevons who, in the late 19th century, wrote that economics was "the mechanics of utility and self interest." One can think of economics as the social science which explores the results of people acting on the basis of self-interest. There is more to man than self-interest, and the other social sciences—such as psychology, sociology, anthropology, and political science—attempt to tell us about those other dimensions of man. As you read further into this book, you will see that the assumption of self-interest, that a person tries to do the best for himself with what he has, underlies virtually all of economic theory.

At the turn of the century, Alfred Marshall's *Principles of Economics* was the most influential textbook in economics. Marshall defined economics as "a study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of wellbeing. Thus it is on one side a study of wealth; and on the other, and more important side, a part of the study of man." Many other books of the period included in their definitions something about the "study of exchange and production." Definitions of this sort emphasize that the topics with which economics is most closely identified concern those processes involved in meeting man's material needs. Economists today do not use these definitions because the boundaries of economics have expanded since Marshall. Economists do more than study exchange and production.

Most contemporary definitions of economics involve the notions of choice and scarcity. Perhaps the earliest of these is by Lionell Robbins in 1935: "Economics is a science which studies human behavior as a relationship between ends and scarce means which have alternative uses." Virtually all textbooks have definitions which are derived from this definition. Though the exact wording differs from author to author, the standard definition is something like this: "Economics is the social science which examines how people choose to use limited or scarce resources in attempting to satisfy their unlimited wants."

Scarcity means that people want more than is available. Scarcity limits us both as individuals and as a society. As individuals, limited income (and time and ability) keep us from doing and having all that we might like. As a society, limited resources (such as manpower, machinery, and natural resources) fix a maximum on the amount of goods and services which can be produced.

Scarcity requires choice. People must choose which of their desires they will satisfy and which they will leave unsatisfied. When we, either as individuals or as a society, choose more of something, scarcity forces us to take less of something else. Economics is sometimes called the study of scarcity because economic activity would not exist if scarcity did not force people to make choices.

When there is scarcity and choice, there are costs. The cost of any choice is the option or options which a person gives up. When you chose to read this book, perhaps you rejected the option of watching television. If you did, the cost of reading this book is the sacrifice of watching television. Most of economics is based on the simple idea that people make choices by comparing the benefits of option A with the benefits of option B (and all other options which are available) and choosing the one with the highest benefit. Alternatively, one can view the cost of choosing option A as the sacrifice involved in rejecting option B, and then say that one chooses option A when the benefits of A outweigh the costs of choosing A (which are the benefits one looses when one rejects option B.)

The widespread use of definitions emphasizing choice and scarcity shows that economists believe that these definitions focus on a central and basic part of the subject. This emphasis on choice represents a relatively recent insight into what economics is all about; the notion of choice is not stressed in older definitions of economics. Sometimes this insight yields rather clever definitions, as in James Buchanan's observation that an economist is one who disagrees with the statement that whatever is worth doing is worth doing well. What Buchanan is noting is that time is scarce because it is limited and there are many things one can do with one's time. If one wants to do all things well, one must devote considerable time to each, and thus must sacrifice other things one could do. Sometimes it is wise to choose to do somethings poorly so that one has more time for other things. (Question 2 at the end of the chapter asks you to illustrate this idea with an example.)

There is, however, an important aspect of economics which conventional definitions of economics ignore. Economists are not interested in examining every case of actions based on costs and benefits, but only on those in which there are some sort of unexpected or unintended consequences. When people live in systems so complex that they cannot fully understand them, they must make choices based on that part of the system which they do know about. These actions can have system-wide implications which are unintended or unexpected by persons making the decision. Economics starts with individuals making choices based on self-interest, but it is primarily interested in how these actions affect society as a whole. Do these choices lead to chaotic results, or to harmonious ones?

When one looks at economics as a study concerned with unintended consequences of human choice and action, one sees that the results people get may not be those that they intend. As was noted above, economists argue that good results do not necessarily come from good intentions, and that good intentions do not necessarily lead to good results. This conclusion has often made economists controversial, and even today there are many who ignore this conclusion. For example, people who try to find a conspiracy behind all the world's problems, whether that conspiracy be of Communists, Jews, bankers, the CIA, or multinational oil companies, start with a belief that bad results must come from bad people with bad intentions.

UNINTENDED CONSEQUENCES AND ECONOMICS

As any other field of study, economics has had a history, and there are books that attempt to trace the trail of economic thought back to its origins. Though the trail can be traced back to the ancient Greeks, it is a difficult trail to follow prior to 1776. In 1776 Adam Smith published *The Wealth of Nations*, a book that is clearly about economics and which inspired a large number of books, pamphlets, and articles in the next 50 years. Before this book most ideas about economics were scattered in writings that were mostly about politics or ethics or philosophy, not in books that were clearly about economics. Yet if one looks at the topics and theories which modern economics textbooks contain and compares them to those things Smith discussed, one is struck by how little of contemporary economics comes directly from Smith. What does come from Smith is a concern about and an interest in unintended consequences.

The most famous term in the *Wealth of Nations* is the term"invisible hand." Smith used this term only once, in the following quotation:

> "...[B]y directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse for the society that it was not part of it."

Another quotation makes clearer what the unintended consequences that the invisible hand leads to are:

> "Every individual is continually exerting himself to find out the most advantageous employment of whatever capital he can command. It is his own advantage, indeed, and not that of the society, which he has in view. But the study of his own advantage naturally, or rather necessarily leads him to prefer that employment which is most advantageous to society."

Smith was not sophisticated in the level of economic theory that he used. He did not understand concepts that are considered basic today, such as the model of supply and demand. (Alfred Marshall developed the modern treatment of supply and demand a century after Smith.) In his comprehensive survey of economic theory, Joseph Schumpeter dismisses Smith as a theorist, saying,

> "The fact is that the *Wealth of Nations* does not contain a single *analytic* idea, principle, or method that was entirely new in 1776."

Though the idea that there could be systematic unintended consequences was "in the air" at the end of the eighteenth century, no one articulated the idea as well as Smith. Because he expressed so well the idea that these unintended consequences are of vital importance for understanding how a society works, Smith has often been called "the father of economics."His book is concerned with a question that has interested economists for two hundred years: under what conditions will actions based on self-interest be beneficial to society. Much of economic theory has been developed and improved in an effort to get better answers to this question.

The two most influential economists in the generation after Adam Smith were David Ricardo and Thomas Malthus. Though they disagreed about many things, they were in general agreement about the topic of population growth, and it was for his writings on this topic that Malthus is best known. Malthus believed that there was a tendency for human population to grow more rapidly than the food supply could be increased. Land was fixed in amount, and more food could be produced either by tilling it more intensively or by adding less-productive land to tillage. In either case an extra hour of labor brought less than an average return of food. The implication of these two different growth rates is clear—eventually a segment of the population will face starvation, and this will cut the growth rate of population.

Malthus' argument on population is pregnant with possible unintended consequences. Suppose that society aids its most needy members by giving them food. As a result, they can survive and reproduce. Helping the poor would, according to Malthus' argument, increase population and in the future lead to even larger numbers on the verge of starvation. Hence, charity would be self-defeating. All attempts to improve society seem doomed to failure according to a strict reading of the Malthusian argument, a truly "dismal" conclusion, as Carlyle noted.

However, Malthus and Ricardo were wrong when they applied this argument to human populations. The predictions they made did not occur. They underestimated both the ability of technology to improve crop yields and the future of birth rates. As it became apparent that they were wrong, economists lost their interest in the study of population. It seemed not to have the unintended consequences for them to explore. Since that time economists have occasionally developed other clever theories with intriguing possibilities of unintended consequences, only to find that they too were in conflict with real-world experience. The reader should bear this in mind as he learns what contemporary economists believe. There is a possibility that today's secure truth may be tomorrow's embarrassing mistake.

The theme that economics is concerned with unintended consequence also helps explain a boundary of economics. The past century has seen the development of the managerial corporation. It may seem surprising that prior to 1870 businesses were not seen as separate entities. A business was identified with its owner. When the owner died or retired, either the business died with him, or it was carried on by members of his family. Today this is no longer true. General Motors is an entity separate from any individual, and it will survive as long as its managers—all of whom can be replaced—are able to keep it alive. This managerial revolution took place at the end of the 19th and at the beginning of the 20th centuries.

What may at first glance seem remarkable is how little this revolution has affected economics. Most choices involving scarce resources and most production in our economy take place in organizations, either in large for-profit organizations or in large not-for-profit organizations. If one considers economics the study of choice or of production, one would expect major changes in economic theory as a result of the managerial revolution. However, it is difficult to find much change in economic theory that is due to the managerial revolution. One explanation of this situation is that economists have not been able to see important unintended consequences that flow from the change. In terms of unintended consequences that affect society at large, that is, the sort of effects that Adam Smith talked about, no one has shown that it makes much difference whether hired managers or legal owners run a firm. Because economics has neglected the large organization, other fields of study have developed to train the managers of these organizations, fields such as business administration, management, marketing, accounting, and finance. Though these fields borrow heavily from economic theory, they are distinct from economics.

On the other hand, even though bureaucratic governments are several thousand years old, the recent growth of government bureaucracy in market economies has affected economics. Not only are there many unintended consequences which flow from government decisions, but government decisions often attempt to change the pattern of unintended consequences which markets produce. Economists are interested in the internal organization of the government and spend much more time studying the management of government than they do with the management of private organizations.

ECONOMICS AS A SCIENCE

Several of the definitions given earlier in this chapter said that economics is a science. Most economists will argue that economics is indeed a science. To examine this claim, it is necessary to begin by defining science.

Philosopher Karl Popper developed a definition of science which has been widely accepted. Popper said that a statement could be called scientific if it is open to the logical possibility of being found false. This definition means that we evaluate the accuracy of a scientific statements by testing it, by comparing it to the world about us. A statement is nonscientific if it takes no risk of being found false, that is, if there can be no way to test the statement against observable facts or events. Popper called this distinction the line of demarcation.

An implication of Popper's definition is that one can never be completely sure that any scientific theory is true. Accepted scientific theory is only theory that has not yet been contradicted by evidence, though the future may bring a contradiction. For example, we cannot be absolutely sure that the statement, "The sun will rise in the east tomorrow" is true because it is a scientific statement. We can easily think of a logical possibility which would refute it—a sun rise in the west. We have great confidence that such an event will not happen because the sun has always risen in the east. However, the fact that all previous experience has been consistent with the statement can never prove that the statement will never be refuted.

Popper saw the growth of scientific knowledge as a process of conjecture and refutation. Someone originally comes up with a way of explaining a set of facts, a conjecture or guess or theory about how the facts are related. If further observation is inconsistent with the theory, the theory is considered refuted and a new theory or conjecture must be found. (Note that if the original explanation is nonscientific, it will never be refuted and there will never be any need to change beliefs. See question 3.)

Most economists see their discipline as scientific in Popper's sense of the word. Economic theory makes statements about how facts fit together, and there are constantly new sets of facts arising that allow one to test the theory to see whether the facts are as theory predicts. However, this process is more difficult for economists than it is for physical scientists.

Unlike physical scientists, economists can almost never use controlled experiments to gather facts with which to test theories. Rather they must use whatever facts the world gives them and then use statistical procedures to try to draw conclusions. Though statistical procedures let economists hold some variables constant to see the effect of other variables, just as a controlled experiment does, they are subject to serious limitations. If there are variables that the theory says are important, but they cannot be measured or they can be measured only imperfectly, statistical procedures may give misleading results. Or the procedures may fail if the theory is uncertain exactly which of the many possible variables that may be involved must be controlled. A strength of a properly done controlled experiment is that there is no need to list all the factors which are controlled. The procedure is such that only one factor, or a small and known group of factors, is different between the control and experimental groups. Given these difficulties, it is not surprising that controversy about whether a theory is supported or rejected by the facts can last for many years in economics.

There is a minority of economists, however, who do not see economics as scientific in Popper's sense. The Austrian school, for example, has argued that economics starts with assumptions and that economic theory is the logically deduced results of those assumptions. If the theory does not fit the facts, one cannot conclude that the theory is wrong, but only that it is inappropriate to apply the theory in that particular situation because the initial conditions do not agree with the assumptions of the theory.

POSITIVE AND NORMATIVE STATEMENTS

Economists make a distinction between **positive** and **normative** that closely parallels Popper's line of demarcation, but which is far older. David Hume explained it well in 1739, and Machievelli used it two centuries earlier in 1515. A positive statement is a statement about *what is* and which contains no indication of approval or disapproval. Notice that a positive statement can be wrong. "The moon is made of green cheese" is incorrect, but it is a positive statement because it is a statement about what exists.

A normative statement expresses a judgement about whether a situation is desirable or undesirable. It expresses a judgement about *what ought to be.* "The world would be a more beautiful place if the moon were made of green cheese" is an example of a normative statement. Notice that there is no way of disproving this statement. If you disagree with it, you have no sure way of convincing someone who believes the statement that he is wrong.

Economists have found it useful to make this distinction because it allows people with very different views about what is desirable to communicate with each other. Libertarians and socialists, Christians and atheists may have very different ideas about what is desirable. When they disagree, they can try to learn if their disagreement stems from different normative views or from different positive views. If their disagreement is on normative grounds, they know that their disagreement lies outside the realm of economics, and economic theory and evidence will not bring them together. However, if their disagreement is on positive grounds, then further discussion, study, and testing may bring them closer together.

Economists can confine themselves to positive statements, but few are willing to do so because such confinement limits what they can say about issues of government policy. Both positive and normative statements must be combined to make a policy statement. One must make a judgement about what goals are desirable (the normative part), and one must decide on a way of attaining those goals (the positive part). Economists often see cases in which people propose courses of action which will never get them to their intended results. If economists limit themselves to evaluating whether or not proposed actions will achieve intended results, they confine themselves to positive analysis. (You should realize that though economists can speak with special authority on positive issues, even the best can be wrong.) However, most economists prefer a wider role in policy analysis, and include normative judgements as well. On normative issues economists cannot speak with special expertise. Put somewhat differently, addressing most normative issues ultimately depends on how one answers the question,"What is the meaning of life?" One does not study economics to answer this question.

Finally, note that most statements are not easily categorized as purely positive or purely normative. Rather they are like tips of an iceberg, with many invisible assumptions hiding below the surface. Suppose, for example, someone says, "The minimum wage is a bad law." Behind that simple statement are assumptions about how to judge whether a law is good or bad, or normative statements, and also beliefs about what the actual effects of the minimum wage law are, or positive statements.

HOW TO APPROACH THE STUDY OF ECONOMICS

Agatha Christie wrote a series of mystery novels in which she challenged the reader to outwit her fictional heroine, Miss Jane Marple. By the end of the book the reader has the same facts which Miss Marple has. But the facts do not speak for themselves. Rather it is Miss Marple's ability to look at those facts in a special way, to see something significant where most readers see nothing, that lets her solve the mystery.

Facts in economics, as in an Agatha Christie novel, need to be organized in some way before they can tell one anything. By themselves they are meaningless. Thus, the study of economics involves more than a memorization of facts. Economics tries to organize facts with its theory. Good theory tells us which facts are important and which are not, and what is cause and what is effect. *The study of economics involves learning how to organize facts the way economists do.*

People who do not understand economics still try to make sense of the world around them by trying to see pattern in the facts they observe. Sometimes they use a simplistic "good-vs-bad" model. In a good-vs-bad model there are two conflicting groups who are classified as good people and bad people. These groups are usually involved in a *zero-sum game*: one person's gain is another's loss. Further, evil motives, possessed by the bad people, lead to bad results unless these people are in some way controlled. Good motives lead to good results.

An example of a good-vs-bad viewpoint was expressed at a town meeting of a small Indiana community during the Winter of 1977. The meeting focused on the natural-gas shortages which the community was facing. One citizen declared that the town faced not an energy problem, but a pricing problem. He noted that several years previously there had been shortages of gasoline at 40 cents a gallon but no shortages at 60 cents. Therefore, he declared, there must have been a conspiracy at that time by oil companies to increase prices as there was now by gas producers. The events he observed do fit into a good-vs-bad framework. He saw a bad result. He saw a bad motive—the desire for profit seems to many people the same as greed or avarice. To connect motive and result, he inferred the existence of a conspiracy.

Though a good-vs-bad model is sometimes appropriate (especially in small-group situations—see Chapter 3), economists are very reluctant to use it. The model of supply and demand developed in Chapters 4 and 6 gives a more sophisticated interpretation of the gasoline shortage, one which is depersonalized and unemotional with no bad groups involved. This model suggests that, in cases of shortages, one should search for government regulation of prices. The good-vs-bad model does not suggest that such regulation is something one should look for. In fact there were price restrictions in place at the time, and Chapter 6 explains why such restrictions could lead to shortages. The good-vs-bad view of the world is attractive because we are able to understand the model at a very young age and because we see the model used so often: in fairy tales, in comic books, in movies, and in television shows, among other places. Because we know how to use this model, and because our culture discourages use of alternative explanations such as fate or mystery, it is easy to fall back to this model if we do not have a more sophisticated model to explain our world.

Economic issues affect us all and most people have opinions about them. These opinions may be based on a good-vs-bad view, some other noneconomic framework, or simply slogans which are often repeated. Often the hardest problem students have in learning about how economists interpret the world is in unlearning their old, noneconomic views.

A well-known example illustrates how hard it can be to unlearn old ways of thinking. In the late 15th century Christopher Columbus believed that by sailing a relatively short distance to the West, he could reach Asia. Contrary to popular myth, it was Columbus, not his critics, who had an outdated view of the world. He believed that the world was much smaller and that Asia was much larger than they actually are; his critics in their guesses were much closer to the truth.

Columbus made four trips to the Caribbean, but he never realized the significance of what he had found. He died believing that he had found a short-cut to the Far East. Rather than use the facts he had before him to alter and improve his ideas of world geography, he insisted on keeping his old views and trying to make the facts fit in.

Economic education involves learning to see reality from new perspectives. Sometimes these new perspectives may surprise you. They may even shock you. But if you take the time to look at the reasoning behind these economic ways of looking at things, you will find that they consist of carefully thought-out and applied common sense.

SUMMARY

This chapter introduced some key ideas of economics such as self-interest, choice, scarcity, and unintended consequences. It also examined the claim that economics is a science and introduced a distinction between positive and normative statements. Finally, it stressed that learning economics is not a matter of memorizing new facts but of learning to see facts in new ways.

QUESTIONS

1. One of the important choices which people make in life is what to do after they leave high school. List as many of the options that were available to you as you can. How many did you seriously consider?

2. What is something you have done that you did not do as well as you could have done? Explain why in terms of costs and benefits. (Comment: This question is *not* asking you to describe a mistake you once made.)

3. a)Which of the following are scientific statements according to Popper's definition? If a statement is scientific, explain what set of events would refute it.

Smoking a pack of cigarettes a day increases one's chance of lung cancer.
Two plus two equal four.
To be a good person, one must be a generous person.
God created the world in 4006 BC.
The Axis powers lost WWII because God was on the side of the Allies.
If the harvest is bad, food prices will rise.
The stars are distant suns.

b) In some societies all unexpected events are explained as manifestations of magic or the intervention of spirits. For example, it the harvest is bad, the gods are angry, and if the harvest is good, the gods ar happy. Explain why this way of interpreting events can persist indefinitely.

4. (a) At the end of the 17th century, England instituted a tax based on the number of windows a person had. Parliament wanted to force those who were wealthier to pay more, but did not have any way to enforce an income tax of the sort we have. At the time, however,

windows were a luxury item, so there was a high correlation between wealth and the number of windows a person had. Hence Parliament taxed windows. What unintended consequence resulted from this tax? (Hint: How could the wealthy avoid the tax?)

(b) Critics of drivers education programs claim that the programs do not reduce the number of accidents that teenagers are involved in. How could this program make teenagers safer drivers and yet increase accidents?

5. Which of the following statements are normative and which are positive?
New York city is on the Hudson River.
Washington D.C. is the capital of the United States.
Washington D.C. is a prettier city than is New York.
Washington D.C. has more people than New York.
When prices go up, people buy less.
When prices go up, poor people cannot buy as much.
High prices are unfair to the poor.

6. Write a statement that is clearly a positive statement. Write one that is clearly normative.

APPENDIX: A NOTE ON MATHEMATICS AND GRAPHS

As you proceed through this book, you will encounter many graphs and some mathematical equations. Economists use graphs and equations because they are quick and clear ways to express ideas. Very often a couple of graphs are worth a thousand words.

There are three sorts of ideas for which graphs and equations are especially well suited. The first is relationships. Examples or relationships from outside of economics are: the higher the altitude, the lower the boiling point of water; the further from the sun a planet lies, the slower its speed in its orbit; the saltier the water, the heavier it is. All of these relationships can be expressed more precisely in the form of graphs and equations than they can be in words. The same will be true of the economic relationships which you will meet in later chapters.

There are two kinds of relationships. If larger values of variable A cause larger values of variable B, the relationship is **direct**. For example, as income rises, people spend more. On the other hand, when larger values of variable A cause smaller values of variable B, the relationship is **inverse** or **indirect**. For example, when the price of apples rise, people will buy fewer.

A second use of graphs and equations is to specify limits. As you now know, economic activity is grounded in the need of choice caused by scarcity. As a result, many economic relations can also be interpreted as limits. You will also see examples of this use of graphs as you read further.

A third use of graphs is to show how a series moves over time. If you look at the back of the *Wall Street Journal* or in the financial section of most large newspapers, you will see examples of this sort of graph.

If you have a good understanding of high school algebra, you will have no trouble at all with the level of mathematics used in this book. You will, for example, realize that algebraic equations can be illustrated graphically and that graphs can be represented algebraically. (Because they are visual and easier for the beginner to grasp, this book relies more heavily on graphs than equations.) If you do not have a good mathematical background, take heart. First, most ideas that yield graphs are first explained verbally and then in tables. And second, the course will give you an opportunity to practice and learn some of the basic mathematics which you should have learned in high school.

QUESTION

1. Illustrate what each of the three relationships mentioned (altitude-boiling point, etc.) above look like. Which are direct and which are inverse relationships?

CHAPTER 2

ACTIONS AND RESULTS

After you read this chapter and work the problems at the end, you should be able to:

Define the following terms: multiplier effect, equilibrium condition, adjustment process, production-possibilities frontier, and budget constraint.

Explain what an identity is and give examples.

Distinguish between an unstable and a stable equilibrium.

Explain why the system of market exchange can be considered as a system of constrained behavior.

Explain the logic behind the problem of the commons.

If given a simple model of contingent behavior in the form of Table 2, be able to determine positions of equilibrium.

People's actions can have surprising results. If they did not, there would be little for economists to study. This chapter considers some situations in which unexpected results can come from interactions of people's actions. In the process of looking at these situations, it introduces you to a number of concepts which are fundamental to the economic way of thinking. The chapter begins by looking at cases in which matching problems constrain the system so that certain results occur regardless of how people behave. It continues by looking at contingent behavior, or situations in which the behavior of each person depends on the expected behavior of others. It then looks at situations in which constraints force people to a result they do not necessarily want in their efforts to get something that they do want. It closes by looking at situations in which an unintended result is a side-effect of other actions.

MUSICAL CHAIRS

In September of 1982, railroad engineers across the United States went on strike for several days until the President signed a bill ordering them back to work. The strike created transportation difficulties in several major cities, including Chicago, as suburbanites who normally rode the train tried to find alternative ways to get to work. Many found that highway congestion was not as serious a problem as was parking-lot congestion. When they finally got downtown there were no parking spaces left. After reporting the problem, one television newscaster in Chicago advised all commuters to start driving earlier the next day so they would find a spot to park.

If all commuters had followed this advice, the situation would have been worse rather than better. The parking lots would have filled up earlier and there would have been the same number of people who would have been unsuccessful in finding spots. Everyone would have wasted time by starting earlier. The situation here is not unlike the proverbial fire in the theater. If everyone calmly walks to the exits, the theater will be emptied as fast as, if not faster than, if everyone runs to the exits. What may be a solution for one is not a solution for all if everyone tries it.

As another example of a matching problem, consider this excerpt from a letter which many thousands of people have received:

> "You too will be making easy money in three to four weeks....My method is simple. I sell thousands of people a product for $5.00 that costs me pennies to produce and mail. Everyone involved in the program makes money from everyone else in the program. If anyone makes money, everyone makes money....This plan has the effect and appeal of a chain letter, but it is legal since you are offering a legitimate product to your people."

The product being sold is a copy of instructions which the buyer then can sell to others.

Is it possible for everyone involved to make money from this scheme? To see that it is not, you must first realize that since the "product" being sold is worthless, the scheme is simply a way of redistributing money. It works with the same logic as a chain letter. All the money that people make in the scheme must come from others involved in the scheme. The sum of money received must equal the sum of money sent. The redistribution is the same as that which takes place in a poker game: the sum of winnings must equal the sum of losings. If anyone involved in the game ends up with *more* than he started with, someone else must end up with *less*.

In matching problems there is no need for any assumptions of how people act; the results are determined regardless of how people behave. Economists do use matching formulas, but never alone as a complete theory. Economic theories always contain assumptions about how people act. But within the theories, matching formulas in the form of **identities** can play important roles. An identity is a statement true by definition. An example of an identity is the statement, "The amount bought must equal the amount sold." Since whatever is bought must have been sold and vice versa, this statement must always be true. Another example of an identity is the statement, "Net worth equals assets minus liabilities." This statement underlies the balance sheet, a basic tool of accounting.

CONTINGENT BEHAVIOR

Contingent behavior exists when each person's actions depend on what he expects others to do. Its role is obvious in our day-to-day behavior. The way a person going to a party dresses will depend on how he expects others to dress. People select the time and place they eat lunch on expectations of when and where others will eat. The route a person takes to work depends on what routes he expects others to use. Students decide whether or not to cut a class before a vacation based on how many other students they expect to cut the class, and the amount they study for a test is often based on how much they expect other people to study.

The existence of unintended consequences in contingent behavior can be illustrated using Figures 1 and 2. In both pictures the behavior of the people involved depends on how they see others acting. You should be able to see that in Figure 1 the end result will be that two people will remain standing. Figure 2 differs from Figure 1 in that one person's instructions have been changed. But in Figure 2 the end result is that all five people remain standing.

TABLE 1

CONTINGENT BEHAVIOR

PERSON		
A		1 other doing so
B	will wear a tie if he sees at least	5 others doing so
C		5 others doing so
D		2 others doing so
E		7 others doing so
F		6 others doing so
G		6 others doing so
H		3 others doing so
I		3 others doing so
J		5 others doing so

There is a **multiplier effect** in these pictures that is the source of the unintended consequence. Changing the instructions of one person changes the behavior of three. When one person changes behavior, it may cause still another to act differently, which in turn can cause still another to act differently. Thus a small initial change may result in a large final change. The result can be unintended because people act ignoring the effect that their behavior will have on the behavior of others.

We could drop the topic of contingent behavior at this point, noting that though it seems commonplace in our day-to-day activity (is it not what peer pressure is all about?), economists have not spent much time with it. However, developing the topic a bit further, by being more specific about the basic model involved in contingent behavior, introduces some ideas that are essential to understanding economic theory. The advantage of meeting them here is that, as with most ideas, they will be easier to understand when you see them in a variety of contexts.

To illustrate the model, suppose there are ten men who make a choice each morning on whether or not to wear a tie to work. Their behavior depends only on how they expect others to act, and this behavior is presented in Table 1.

Based on this information, how many workers will wear ties? To begin answering this, we need to specify an **equilibrium condition**. An equilibrium exists when the system is at rest and has no further tendency to change. The logical equilibrium condition in this model is that there will be no further changes in behavior when expected behavior is the same as actual behavior. When actual behavior is different from expected behavior, we expect some adjustments to be made so the system will not be at rest. To determine

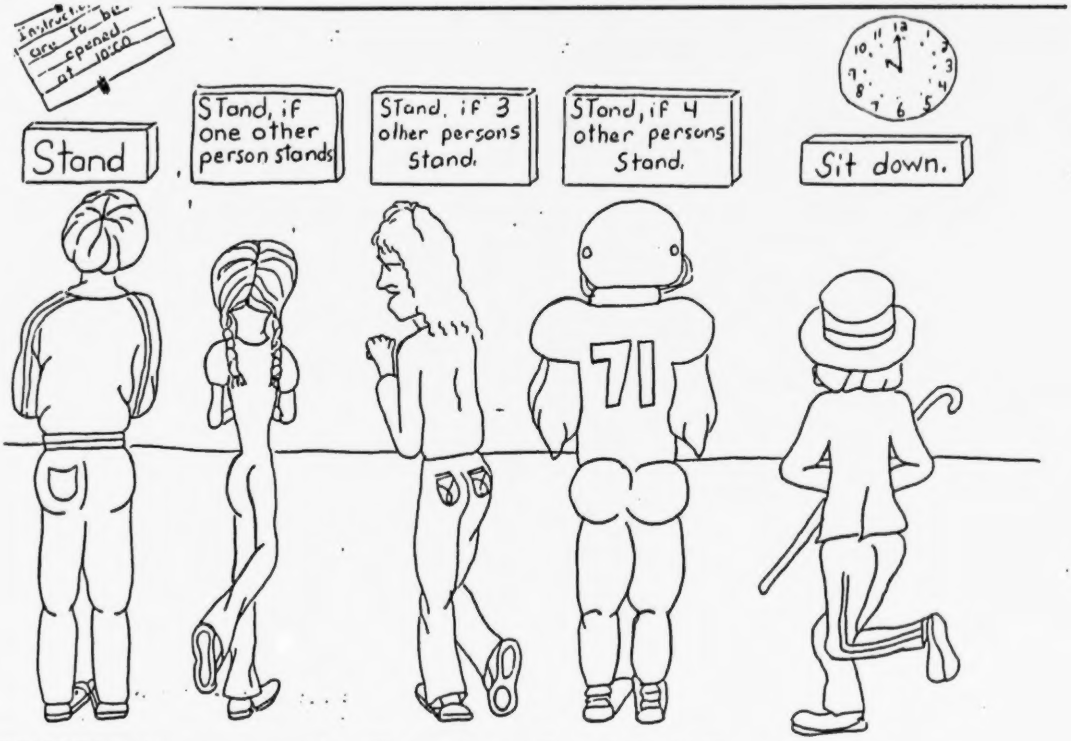

Instructions are to be opened at 10:00
Stand
STand, if one other person stands
STand, if 3 other persons Stand.
STand, if 4 other persons Stand.
Sit down.
71

FIGURE2.1 How many people will be left standing after 10:00?

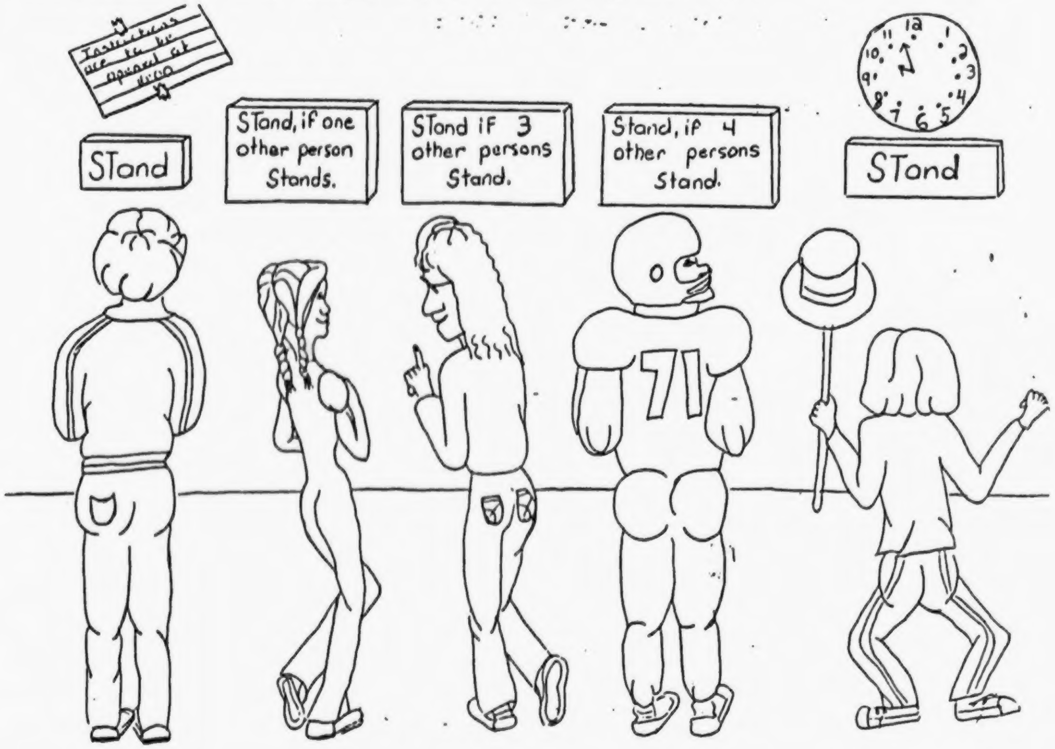

Instructions are to be opened at 10:00
STand
STand, if one other person Stands.
STand if 3 other persons Stand.
Stand, if 4 other persons Stand.
STand
71

FIGURE2.2 How many people will be left standing after 10:00? You should see the multiplier concept at work here. Explain how it works in this case.

where the system will be in equilibrium, the information in Table 1 can be reorganized to yield Table 2. Table 2 says that if no one is expected to wear a tie, no one will. If one other person is expected to wear a tie, one person (A) will. If at least 5 others are expected to wear a tie, everyone except E, F, and G, or seven people, will wear ties.

By trying various possibilities, one will discover that there are three positions of equilibrium: no one wearing ties, everyone wearing ties, and four people wearing ties. This is done in the third and fourth columns of Table 2. Thus, if five other people are expected to wear a tie, seven in fact will. All the people who wore ties will be satisfied, because at least five other people were wearing a tie. But persons F and G will not be happy. They will observe that seven other people are wearing ties, and will change their expectations and thus their behavior.

This model has several parts. One part, given in Table 1, is a statement or assumption about how people act. As in most economic theories, this model makes assumptions about how *individuals* act, but is interested in group results. A second part of the model is a condition of equilibrium. Implicit in our discussion above is a third part, an **adjustment process.** An adjustment process explains how a system moves from a position of disequilibrium to one of equilibrium. An important part of the adjustment process in this model would be statements of how expectations were formed. The simplest would be that people expect others to continue acting as they did the day before. You can act out what happens with this assumption. If ten students are given cards with instructions similar to those in Table 1 (instead of wearing ties, the instructions might say, "Raise your hand." or "Sit down."), a class can experiment with various starting points and see what happens.

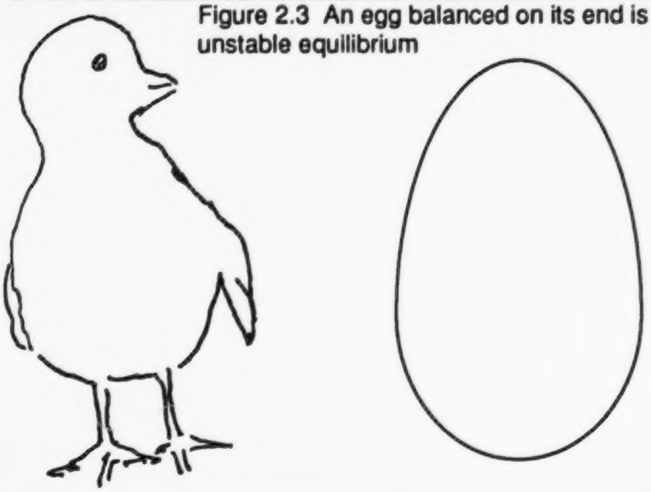
If four people wear ties, the system will be in equilibrium, but the equilibrium will be **unstable**. This means that any small divergence away from equilibrium will take the system increasingly far from it. If the system begins with an equilibrium of four, and A forgets his tie one day, the system will move to zero. If E wears a tie for a special reason, the system will move to ten. In contrast, zero and ten are **stable** equilibria. (The plural of equilibrium is equilibria. The word thus betrays its Latin origin.) A small divergence from either will tend to disappear. Figure 3 illustrates an unstable equilibrium. An egg balanced on its end is at rest with no tendency to move. Hence, it is in equilibrium. But if anything disturbs it, it will not return to this original position. Filling the bottom of the egg with plaster would make this equilibrium stable. Then any disturbance would rock the egg, but it would return to its original position.

Models of contingent behavior can be more complex than that illustrated in Table 2.

TABLE 2 CONTINGENT BEHAVIOR REVISED

OTHERS EXPECTED	NUMBER ACTING	WEARERS OK?	NON-WEARERS OK?
0	0	—	YES
1	1	NO	YES
2	2	NO	YES
3	4	YES	YES
4	4	NO	YES
5	7	YES	NO
6	9	YES	NO
7	10	YES	—
8	10	YES	—
9	10	YES	—

Figure 2.3 An egg balanced on its end is an example of an unstable equilibrium



What if, for example, not everyone has the same initial expectations or if they adjust to new expectations in a variety of ways? Further, the model may not be good for situations involving small groups, such as we used for illustration. In a small group people will not only respond to and form expectations on the basis of observed behavior, but they also will rely on verbal conversations among group members.

There are a number of economic situations that contingent behavior may help explain. The prices of fine art, stamps, rare coins and other collectables seem in large part determined by what people expect others to be willing to pay. People are willing to spend $5000 or more (and often much more) for a postage stamp because they are confident that they can resell it at the same price or a price close to it. Contingent behavior also may explain the wide price fluctuations in some speculative markets, such as that for gold. It even explains why people consider paper money valuable. Everyone is confident that others will accept it.

However, economists have not spent much effort exploring contingent behavior with one exception. The one exception involves the Keynesian theory of national income, which was for many years the primary way economists thought about how unemployment and GNP were determined. In this theory people's spending depends on what they expect their income to be, and income in turn is the same as spending (because whatever one person spends, another receives as income). Equilibrium in this model exists when people receive the income they expect—otherwise they change expectations and behavior. There is also a multiplier effect in this model. When one person changes his behavior, it affects others, who change their behavior in turn. (Chapter 12 in volume 2 develops this model.)

CONSTRAINED BEHAVIOR

A probable reason economists ignore models of contingent behavior is that these models do not clearly present the favorite assumption of economists, that individuals base their actions on the costs and benefits which they see. The importance of benefits in determining behavior implies that people have goals, for without goals, one cannot talk about benefits of a decision. Costs exist whenever there is scarcity and choice. Cost is whatever options must be given up in making a choice. Often cost is in the form of money, but it can take other forms as well. The cost of going to college, for example, includes not only the money a person spends on tuition, (which could be spent on something else) but also includes the time spent studying and going to classes. The value of this time can be estimated by computing the amount of income a person could have earned if he did not go to college.

Scarcity is at the basis of many economic concepts because it constrains or limits our behavior. Let us explore the notion of constrained behavior by starting with the simplest sort of economic structure, Robinson Crusoe alone on his island.

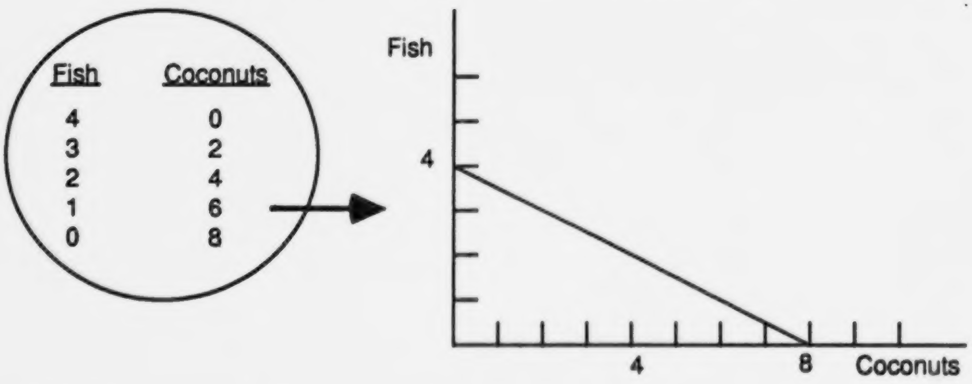
Suppose that each day Crusoe has enough time to catch four fish or to gather eight coconuts. Notice that he cannot have both four fish and eight coconuts. If he uses his time to catch fish, he does not have that time to find coconuts. If he wants both coconuts and fish, he can devote some time to both. If, for example, he spends half the day fishing and the other half gathering coconuts, he can have two fish and four coconuts.

A list of all the possible combinations of fish and coconuts open to Crusoe makes up his **production-possibilities frontier**. The production-possibilities frontier separates outcomes which are possible for an individual (or a group) to produce from those which cannot be produced. Because Crusoe cannot exchange, his production-possibilities frontier is also his **consumption-possibilities frontier**. The consumption-possibilities frontier, which later chapters will also call a **budget constraint**, is the line indicating which outcomes are affordable and which are not affordable. Figure 4 illustrates Crusoe's production-possibilities frontier (and his consumption-possibilities frontier).

The slope of the frontier in Figure 4 indicates the costs facing Crusoe. In order to get an extra fish, he must sacrifice two coconuts, and to get another coconut, he must sacrifice one half a fish. Notice that there is no money involved; cost does not depend on money. In economics the cost of anything refers to whatever is given up in order to get that thing.

Now suppose that civilization arrives in the form of the Friday Trading Post. The trading post trades one fish for one coconut (or one coconut for one fish). Will the existence of the trading post make life better for Crusoe? It will if it gives him more consumption possibilities than his production-possibilities frontier allows him. To see that it does, consider what will happen if Crusoe collects eight coconuts, and trades them for fish. He will be able to get eight fish, which is not a possibility that his original budget constraint allowed him. Figure 5 shows the consumption-possibilities frontier which exchange opens up to him.

Figure 2.4 Crusoe's Production-Possibilities Frontier



Suppose that before trade, Crusoe was at point a on his production-possibilities frontier in Figure 5, and that once trade is possible, he wants to be at point b. In order to get to b, he must devote all his time to gathering coconuts. Some of those coconuts he will want, but others he gathers only to trade. In order to get to where he wants to be, Crusoe must do something he does not want to do. He must work gathering coconuts.

Exchange, then, requires a person to give in order to get. It makes people do things which they do not want to do in order to accomplish their goals. Adam Smith wrote about constrained behavior in this way:

> "It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinners but from their regard to their own interest. We address ourselves not to their humanity but to their self love, and never talk to them of our own necessities but of their advantages."

The butcher and baker want to profit from their businesses so that they can obtain material goods for their own use. In the pursuit of this goal, they provide a product to the consumer. If one of these merchants refused to provide a quality product at a reasonable price, consumers could take their money elsewhere. The existence of other sellers provides a limit on each merchant. However, they can at times find ways to "beat the system." They can stop competing and monopolize, charging consumers more in the process. Or they can sell defective merchandise to consumers who think that they are buying quality merchandise.

The system of market exchange is the most studied system of constrained behavior with which we live. There are, however, many other such systems around us. Politicians would be less sensitive to people's desires if they did not have to endure elections. Mothers and fathers who tell their children that they will not get dessert if they do not finish their meal are using this system, as are grade school teachers who tell their students that they will get no recess if they are not quiet. Finally, college students deal with an interesting system of constraint called the "grading system" in their struggle to earn a college degree.

Let us consider how the grading system works under some extreme assumptions. Assume that students want to obtain a college degree but are totally uninterested in learning anything in the process of obtaining that degree. Assume also that the faculty has a goal of increasing student knowledge, and does not care at all whether or not the student does graduate.

The faculty can achieve its goals with the

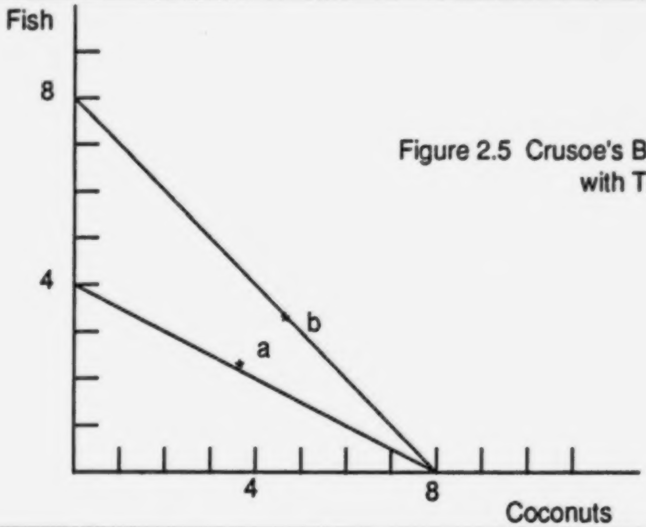


Figure 2.5 Crusoe's Budget Constraint with Trade

system of tests and grades that is familiar to all college students. This system forces students to learn what the faculty members desire in order to obtain a degree. Whether students are fascinated with what they learn or have no interest in it at all does not matter in this system. In either case, learning must take place.

Or must it? College students have long known that there is in fact a way to "beat the system," that is, one can achieve the goal without performing as the system requires. Students can cheat. When they cheat, students still achieve their goals, but they no longer do those things which they do not want to do. All systems of constrained behavior have ways of "cheating" on the system, and if cheating is too easy, the system will not work well.

If one views college education as a struggle to learn, with grades and degrees as unimportant by-products, cheating should never happen. The fact that cheating does occur in colleges suggests that there is a large element of constrained behavior in college education.

Sometimes systems of constrained behavior have been consciously designed, but often they have grown up spontaneously, as did the market system. Whether spontaneous or designed, people often try to change how systems work. When a system of constrained behavior is complex, and those making the changes do not fully understand the system, the results of the changes in behavior can seem surprising and unexpected to those making the changes. For example, consider the grading system. If the faculty believed students are motivated only by a love of learning, it might decide to abolish grades. It would be surprised when the amount of learning falls. When one wants to alter a system of constrained behavior, it is necessary to understand the system. This necessity provides the major demand for economists. Governments often want or need to change incentives in markets, and they often ask economists for predictions about the effects of those changes. And even when they do not ask, economists give answers. A great deal of their scholarly work is concerned with how the system of market exchange works, and what the effects of various changes will be. You will see many examples of economists concern for the effects of such changes in the chapters ahead. In some of these examples you will see how useful economists find the simple concepts of budget constraint and production-possibilities frontier.

BY-PRODUCTS OR SIDE-EFFECTS

Sometimes when a person tries to achieve his goals, side-effects occur which affect others whom the person making the decisions either does not know about or does not care enough about to let them influence his decisions. Economists have found a number of situations which are based on such side-effects. An obvious example of by-products is the production of pollution. However, by-products may be helpful or neutral to others as well as harmful.

Pollution is not an intended product, but it is one output in a production (or consumption) process. Consider, for example, the production of electricity from coal. Among the outputs of this production process are electricity, heat, carbon dioxide, and ash. The type of coal used will influence the outputs—some coal is high in sulfur, which will emerge in some form as an output. The electricity is the desired output, and the others are unwanted outputs that cannot be totally avoided. If one of the outputs is considered undesirable, it can sometimes be controlled, but this process will create other by-products.

Pollution is a word that refers to by-products that are considered undesirable to others. If no one considers the waste heat that is a by-product of the production of electricity a problem, that waste heat will not be called pollution. When it is considered a problem, the government may force the electric utility to use resources to deal with it, by building cooling towers, for example. A cooling tower does not, of course, reduce the total waste heat. It only transfers some of it from the water to the air where it is considered less undesirable.

One of the most important topics in the class of by-products is the creation of money. Money in modern economies arises as an unintended results of banks in the process of making loans. This aspect of banking is important to economists because they have evidence that changes in the amount of money influences the amount of inflation and unemployment which a society has.

An equally important case of actions with a by-product is one which has become known as the **problem of the commons**. The commons exist whenever there is a resource, such as grazing land, a hunting area, or a fishing grounds, which is open to all. The problem of the commons refers to the absence of any automatic mechanism or incentive to prevent the overuse and depletion of the commonly-held resource.

To see why this overuse can happen, consider a common grazing area. Suppose Table 3 shows the relationship between the number of cows using the land and the amount of milk that can be produced. (We shall ignore all other outputs, such as meat, leather, and fuel. The technical name for a table such as Table 3 which shows output depending on inputs is "production function.")

If one person controlled access to this land, it would be foolish for him to allow more than 400 cows to graze there. But when the land is open to all, 500 cows may end up grazing if each of the many herders tries to do what is best for himself. Thus, if there are 100 herders, each with four cows, each cow will produce 7 gallons of milk. If one herder adds an extra cow, the total production of milk does not increase, *but the production from his herd will.* He will have five cows, each producing 6.98 gallons (ignoring all digits beyond the second to the right of the decimal point). The fifth cow adds 6.98 gallons of milk. However, each of the four cows he already has will produce .02 gallons less, so the net gain to the herder will be 6.9 gallons. He will not take into consideration the drop in production of .02 gallons from each of the 396 cows that he does not own. Since each herder has the same set of incentives, it may easily happen that there will be 500 cows grazing on the land. Clearly there is an unintended consequence here, and it is of the by-product type. Each herder ignores the effects that his cows have on the productivity of cows he does not own.

TABLE 3
PRODUCTION ON
THE COMMONS

# OF COWS	OUTPUT OF MILK
100	1000 GAL
200	2000
300	2600
399	2799
400	2800
401	2800
500	2700
600	2600

The logic of the problem of the commons is behind many ecological problems, from deforestation to depletion of the ozone layer. In a sense the core of the problem is that behavior is unconstrained in any way. The next chapter will describe two basic methods that restrictions can be placed on people's actions in order to give better results.

SUMMARY

In exploring some of the reasons for unintended or unexpected consequences, this chapter introduced you to a number of important concepts and ideas which you will need to know as you read further. You met concepts of identity, model, behavioral statement, equilibrium condition, stable and unstable equilibria, adjustment process, budget constraint, and production-possibilities frontier. You also met the assumption that people act on the basis of cost and benefits, an assumption which you will see throughout the book. Finally, you saw examples of constrained behavior, by-products, and the problem of the commons.

QUESTIONS

1. "In Ithaca, New York, where I live, the cable TV system carries most New York Yankee baseball games. One August night, sportscasters Phil Rizzuto and Bobby Murcer were calling a slow game between the Yankees and the Milwaukee Brewers. Between pitches, Rizzuto was looking over his record sheets and remarked that the Brewers had done much better in day games than in games played at night. Murcer checked his own records and found that the Yankees, too, had a much higher winning percentage during the day. With characteristic enthusiasm, Rizzuto then conjectured that all teams have better records for day games. In a brisk exchange of the sort that makes summer evenings in Ithaca seem to fly by, the two then spent the rest of the inning discussing the poor lighting conditions in American League parks and various other difficulties that might help account for why teams do so poorly at night."

(From Robert H. Frank, *Choosing the Right Pond: Human Behavior and the Quest for Status* (New York: Oxford University Press, 1985), p. 3.)

There is something terribly wrong in the discussion described in the previous paragraph. What is it and why is it wrong?

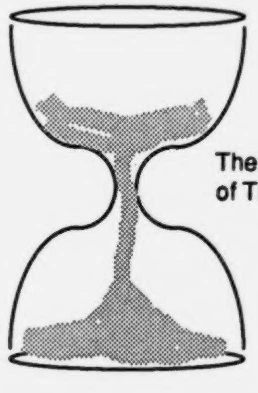
2. Replace the numbers in Table 1 with the numbers below. What will the equilibrium (or equilibria) be? Is it (are they) stable?

A will wear a tie if at least 7 others do.
B will wear a tie if at least 9 others do.
C will always wear a tie.
D will never wear a tie.
E will wear a tie if at least 6 others do.
F will wear a tie if at least 3 others do.
G will wear a tie if at least 5 others do.
H will wear a tie if at least 2 others do.
I will always wear a tie
J will never wear a tie.

3. When is the hourglass at the right at equilibrium? Is equilibrium always a good thing? Explain.

4. Suppose that a cure for all cancer is discovered. Why will this increase the percentage of people who die from heart disease? What idea in this chapter does this possibility illustrate?

5. Which would an economist expect to be more likely to reduce cheating: repeatedly telling students that cheating harms other students and is immoral behavior, or increasing efforts to detect cheating and increasing the penalties?


The Sands of Time

6. a) Using Table 1 in the chapter, suppose that initially persons E, F, and G are wearing ties. What will the equilibrium be? Describe the steps in which the system reaches equilibrium.

b) Suppose that initially people B, C, G, F, and J are wearing ties. Describe the adjustment process.

c) Suppose person E has the instructions, "Always wear a tie." What will this change do the positions of equilibrium?

7. Which salesman would you expect to be more likely to cheat his customers: a door-to-door salesman, or a salesman from a large steel company selling to a major automobile manufacturer? Explain your answer in terms of the costs and benefits of cheating.

8. Suppose that Crusoe and Friday live together on an isolated island. Each day Crusoe can either catch four fish or gather eight coconuts, and Friday can either catch six fish or gather 3 coconuts. Compute the production outcomes of this economy in the table below, and then explain which one of the four possibilities will not be on the production-possibilities frontier.

	total coconuts	total fish
Both fish		
Friday fishes, Crusoe gathers		
Crusoe fishes, Friday gathers		
Both gather		

9. The Dukedom of Scimonoce (it is a backward economy, hence its name) is a totally agricultural economy that grows only corn and/or oats. There is no trade with neighboring countries, and there are only five small plots of farmland that produce as follows:

Plot	either	Corn	or	Oats
A		50		10
B		40		10
C		30		10
D		20		10
E		10		10

a) If the country decides that it wants 30 units of oats produced, what is the largest amount of corn that can be produced? ___________ Which plots do you plant in oats to get this result, and which do you plant in corn?

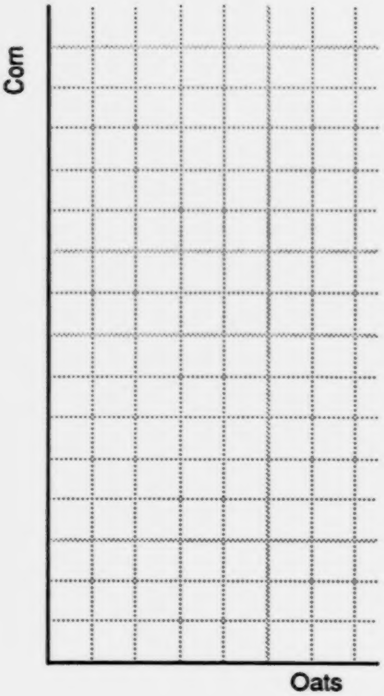
In Corn _______

In Oats_______

b) Given your answer in part (a), what is the opportunity cost of 10 more units of oats (increasing oat production from 30 to 40)?_____________

c) Finish filling in the production possibilities table below and plot the points on the graph.

CORN	OATS	Plots Used for Corn	Plots Used for Oats
150	0	A, B, C, D, E	None
140	10		



d) Where is the point 30 oats and 30 corn on the graph? (Label it X.) Is it possible to produce 30 units of corn and 30 units of oats? If you answer that 30 corn and 30 oats is attainable, explain how you get it, that is, where do you plant oats and where do you plant corn.

e) Is the production of 30 corn and 30 oats efficient?

f) Where is the point 40 oats and 60 corn on the graph? (Label it Y.) Is it possible to produce 40 oats and 60 corn? If you answer that 40 oats and 60 corn is attainable, explain your planting decision that makes it possible.

g) Will the dukedom produce up to its potential if plot A is left empty? What will happen to the production-*possibilities* frontier?

h) Suppose consumers decide that they want 1000 units of corn and 300 units of oats. Will anything happen to the production-possibilities frontier? Explain.

i) Suppose the owner of plot E finds a method to increase his (and only his) oat production to 20. Recompute the production-possibilities table and draw the new curve on the graph above.

j) From the information given, can one tell which combination of oats and corn the dukedom should produce? Explain.

10. "I had long been puzzled by the behavior of the whaling industry.... [W]hales, which could provide a permanent resource if harvested prudently, were overexploited and driven toward extinction. This mysteriously "self-destructive" behavior of the industry was explained to me by a Japanese economist. Like most conservationists, I had assumed the industry's goals would be to maximize the sustainable annual yield of whales. In fact the industry has been trying to maximize the present value of the whale resource. If various species of whales are driven to extinction... the capital can be turned to the rape of another (in the minds of economists) array of available resources. Such behavior [is] perfectly rational in the dominant economic paradigm...." (From Paul Erlich, "Environmental Disruption: Implications for the Social Sciences," *Social Science*

Quarterly, March 1981, p 10.) From what you have learned in this chapter, do you think Erlich (or the unnamed Japanese economist) understands the usual view of economists about why whales are threatened with extinction? Explain. (Hint: Does the whaling industry have goals, or do individual whalers have goals?)

CHAPTER 3

THE INDIVIDUAL AND THE GROUP

After you read this chapter and work the problems at the end, you should be able to: . . .

Explain the logic of the prisoners' dilemma.

Explain how Hobbes differed from Smith in his predictions of the results of actions based on self-interest.

Explain why the assumptions and methods of economics do not allow it to examine the inner workings of families or other small groups who know each other well, but do allow it to examine exchange and politics.

Distinguish between restricted and generalized altruism, and explain why economists can live with the first but not the second.

List the two basic ways in which self-interested decisions can be coordinated for the good of the group.

Economic theory is based on the assumption that individuals pursue their self-interest. Each person makes decisions based on the costs and benefits which that person perceives. The last chapter showed that the pursuit of an individual's goals could affect the group in ways which the individuals could not foresee.

This chapter explores the topics of self-interest, group interest, and the relationship between them. It uses the prisoner's dilemma, which is an example of a technique called *game theory*, to illustrate possible relationships between the individual and the group. Game theory is an analysis of interactions where the outcome a person faces depends not only on his strategy of action, but also on the strategies of others. Since its development in the 1940s, game theory has found important uses in many fields, such as economics, political science, psychology, and biology.

THE PRISONERS' DILEMMA: TWO STORIES

The prisoners' dilemma is the story of two criminals who have been arrested for a heinous crime and are being interrogated separately. Each knows that if neither of them talks, the case against them is weak and they will be convicted and punished for lesser charges. If this happens, each will get one year in prison. If both confess, each will get 20 years in prison. If only one confesses and testifies against the other, the one who did not cooperate with the police will get a life sentence and the one who did cooperate will get parole. Table 1 illustrates the structure of payoffs.

Given this set of payoffs, there is a strong tendency for each to confess. This can be seen by considering the choices and payoffs of either one. If prisoner A remains silent, prisoner B is better off confessing (because parole is better than a year in jail). However, B is also better off confessing if A confesses

TABLE 1 THE PRISONERS' DILEMMA

		PRISONER B's STRATEGIES	
		DO NOT CONFESS	CONFESS
PRISONER A's STRATEGIES	DO NOT CONFESS	1 Year 1 Year	Parole Life
	CONFESS	Life Parole	20 Years 20 Years

TABLE 2 THE PRISONER'S DILEMMA RESOLVED

		PRISONER B's STRATEGIES	
		DO NOT CONFESS	CONFESS
PRISONER A's STRATEGIES	DO NOT CONFESS	1 Year 1 Year	Parole and Death Life
	CONFESS	Life Parole and Death	20 Years and Death 20 Years and Death

(because 20 years is better than life). Hence, B will tend to confess regardless of what A will do, and by an identical argument, A will also tend to confess.

The prisoners' dilemma is a case in which actions determined by self-interest are not in the group's interest (where the group is defined to include only the criminals, not the larger society). It is a story which would have pleased Thomas Hobbes, a political theorist of the mid 17th century. Hobbes is a grandparent of economics because he introduced into intellectual discussion the assumptions that the individual was the starting point of social analysis, and that the individual was motivated by self-interest. He believed that the unrestrained pursuit of self-interest would result in chaos. The development of government, with its power to coerce people, was necessary to bring order out of chaos.

However, the story can be retold with a slight alteration. Assume that our two criminals belong to an organization of criminals which will kill any member who confesses to the police or testifies against other members of the organization. With this change, the outcome of the story will tend to be totally different: both will remain silent. Self-interest and group interest are identical. This happens not because the motives or goals of the criminals have changed, but because incentives or payoffs have changed. Table 2 shows the new payoffs. A paradox of this case is that imposition of this "death penalty" can actually improve the well-being of individuals making up the group. One would not normally think that making possible outcomes worse could improve results.

For Hobbes, the introduction of government could change payoffs so that cooperative behavior would replace the "war of all against all." But Adam Smith better illustrates the results of this second story. His *Wealth of Nations* argues that the well-being of the group can be consistent with the pursuit of self-interest if the right incentives, which Smith said were free trade and competition, existed. Nowhere does he explain his point better when he talks about the butcher, baker, and brewer who provide quality products because they love themselves, not because they love humanity.

If the individuals in a group are motivated by self-interest, the group cannot survive and prosper unless it finds ways to make individual self-interest consistent with group interest. This is a problem for the largest of organized groups, a nation, as well as for many of the subgroups within it. There are two basic ways to deal with this problem. One is the Hobbesian way of regulation and planning from the top. The other is the Smithian way of property rights and the market. Before we consider these two ways in more detail, let us consider some of the assumptions used in the prisoner's dilemma.

THE ASSUMPTION OF SELF-INTEREST

The prisoners' dilemma assumes that the players are selfish, that is, they do not care what happens to the other player. But this assumption is not always true. Sometimes people do worry about others. Suppose, for example, that the two prisoners are fellow revolutionaries dedicated to the ideals of their cause or are husband and wife. In this case, even with the payoffs given in the first table, it is unlikely that either prisoner will confess. Each prisoner in this case considers not just what will happen to him as a result of his decision, but also what will happen to the other prisoner.

However, one should not argue that "good" motives necessarily lead to good results. Suppose there is hard evidence that links two criminals to a crime and that each cares a great deal about the other. If neither confesses, both may serve ten years. Under these circumstances each may confess to the crime, claiming that he alone was responsible in an effort to help his fellow prisoner. But if both try to help the other by admitting guilt, they both may end up serving more time than if each had maintained silence. Of course this would not happen if they could communicate and plan their strategies, but often in life both information and communication are lacking. "Good" motives with poor knowledge can result in harm to the group welfare.

If one ignores problems of information and knowledge, a conclusion one can draw from this chapter's stories about prisoners is that there are two ways for groups (including that largest group, the nation) to get good results. Groups can take motivations or goals as given and try to find incentives (payoffs) that yield desired results, or groups can take incentives as given and try to find motivations or goals which yield desired results. However, a closer look at the assumption of self-interest suggests that the problems of information and knowledge cannot be ignored, and that they place important limitations on the ways groups can organize.

In discussing self-interest, one must distinguish between people's actions and motives in small-group situations and in large-group situations. In small-group situations, as in a family or tribe or village in preindustrial societies, self-interest seems to be only one of many motives determining actions. At least as important are motives such as loyalty, duty, affection, and compassion. In such situations narrow self-interest—the ignoring of the interests of others—is a vice. Small groups often depend on selfless motives to get good results. The most basic of groups, the family, provides a good example. It could not exist without selfless motives. In a family good intentions usually lead to good results because knowledge that family members have about each other is extensive. In small groups it is common for people to consider how their actions affect others in the group.

However, it is dangerous to assume that large groups can be organized in the same way as small groups. As groups get larger, more and more knowledge is needed if good motives are to lead to good results, and less and less is generally available. In addition, as groups grow larger, those affected by a person's actions change from identifiable people to anonymous strangers and this can affect behavior.

When one makes a distinction between small groups and large groups, one finds that the assumptions needed to get "self-interested" behavior are extremely weak. In fact it is possible to assume perfectly altruistic people, people who care only about the well-being of others, and still end up with self-interested behavior. One need only assume that the altruism is very strong toward the small groups with which people are associated (such as families) and weak or nonexistent toward strangers (the large-group world). For example, a man who is interested in the well-being of his family acts in a way that, as far as the large group is concerned, may not be distinguishable from actions motivated by narrow self-interest. He will demand as much pay as he can get for his work and will try to pay the lowest price he can for the goods he purchases. An appeal made to him to work at a lower pay for the common good is likely to be resisted because it will harm the group he is most interested in helping, his family.

Small-group motives sometimes appear in places where large-group motives would seem more appropriate. For example, in discussing the structure of the Japanese corporation, a manager states: "The most important path to quality is that employees feel that they are of one family, playing an important role in the company....In Japan, an employee tries to work his lifetime with the company he chooses. So he does not put much importance in his personal interests....In Japan, they say, 'Job first, family life comes second.'"[1]

Nevertheless, the distinction between small group-motives and large-group motives suggests that it is very difficult to organize large groups on any other basis than on self-interest. This distinction tells us that it is not enough to convert people from selfishness to altruism, but that the altruism they are converted to must be a *generalized* altruism. People must care as much about strangers as they care about themselves, their families, and their friends.

The Soviet Union has a hero who illustrates the sort of change the transformation requires. Pavlik Morozov turned his father over to authorities in 1932 as an enemy of the people. Apparently his father, the mayor of Garasimovka, gave or sold safe conduct passes to peasants fleeing famine. When thirteen-year-old Pavlik learned of his father's crimes and informed the secret police, his father was executed. His uncle and grandfather then revealed their reactionary natures when they retaliated by killing him. Pavlik became a martyr for the glory of the Soviet state and, as a hero, many schools and youth groups bear his name. For a nation to be organized not on self-interest but on "selfless" motives, people must be like Pavlik. They must consider the interests of anonymous strangers to be at least as important as their own interests and the interests of their family and friends.

Economists have been content to take goals as given and have not tried to develop theories of goal formation. This attitude reflects the hypothesis that it is impossible on a large scale to change people so that they act on the basis of generalized altruism. Not everyone accepts this hypothesis. Most Marxist's governments, governments which invoke the name of Karl Marx as a justification of their policies, believe or profess to believe that man can be transformed so that he does not pursue his own selfish interests but strives only for the welfare of the group. Some of these governments have tried to refute the economists' hypothesis by transforming their citizens. China's Great Leap Forward and Cultural Revolution and Cambodia's bloodbath under the Khmer Rouge are dramatic examples of attempts to make this transformation. Both failed disastrously. If, in the future, someone does refute this hypothesis, it will cause a revolution in economics and greatly diminish the importance of what it has to say. (Needless to say, most economists do not expect this hypothesis to ever be refuted, or else they would not have devoted so much of their time to economics.)

---

[1] *The Wall Street Journal*, April 29, 1983, p. 14.

TABLE 3 THE PRISONERS' DILEMMA REVISITED

Ingrid's Choices

Edward's Choices		Work Hard	Loaf
	Work Hard	1 Hr Work b 1 Hr Work	a Do 1/2 Hours Work in 1 1/2 Hrs Do 1 1/2 Hrs Work
	Loaf	Do 1 1/2 Hrs Work d Do 1/2 Hrs Work in 1 1/2 Hrs	Do 1 Hrs Work In 3 Hrs c Do 1 Hrs Work in 3 Hrs

OTHER ASSUMPTIONS

Our story of the prisoners' dilemma assumed that the prisoners' interacted only once. Since no further interaction will take place, neither prisoner needs to take into account any future consequences of his decision. But in real life many situations which appear to resemble the prisoners' dilemma are situations which will repeat.

Consider, for example, the case of Ingrid and Edward in Table 3. Their mother requires them to clean the house each Saturday before they can play. Each has a choice of working hard or loafing. If they both work hard, they can finish in one hour. If they both loaf, the work takes three hours. If one loafs and the other works hard, the job will take an hour and a half. Ingrid ranks the possibilities in Table 3 from best to worst as *a*, *b*, *c*, and *d*. Edward ranks them in a similar way.

The logic of the prisoners' dilemma suggests that both will loaf. However, because Edward and Ingrid must interact each week, the situation is more complex. One's actions now may be able to influence the other's actions in the future. Each must see their strategy not just in terms of how it affects the payoff this week, but how it will affect the payoffs in the future as well. What one player does will affect the strategy of the other, and in turn this result can affect the strategy of the original player. This chain of decisions coming back to affect future decisions is a case of feedback, and it greatly complicates the problem of deciding strategy.

What strategy should a player choose? There is no one best strategy for all circumstances, but one which works extremely well over a wide variety of environments is a simple tit-for-tat strategy. In this strategy one begins by cooperating, and then mimics the other player's moves. Tit-for-tat is "nice" in that it is willing to cooperate, and it does not bear a grudge. It also cannot be exploited, because any "defection" from cooperation will be returned.

If Ingrid uses a tit-for-tat strategy in our example above, she will work hard the first week. If she then discovers that Edward has loafed, she will loaf the next week. If, after he realizes that he cannot exploit her, he becomes ready to cooperate, a tit-for-tat strategy is ready to begin cooperating.

The power of tit-for-tat in encouraging

TABLE 4 HAND OR FIST

		MAJORITY OF CLASS	
		Hand	Fist
YOU	Hand	20	0
	Fist	30	10

cooperation in unusual places has been explored by Robert Axelrod in *The Evolution of Cooperation* (Basic Books, 1984). He argues that it could explain the many spontaneous outbreaks of "peace" in trench warfare during World War I. In the early stages of this warfare, informal and unspoken agreements often existed in which neither side would fire to kill the enemy when major battles were not underway. The best position for the soldiers on each side would have been one in which they could "exploit" the enemy by killing him with no fear of retaliation. Then when a major battle came, the enemy would be weaker and the chances of surviving the major battle would be better. However this was not an attainable position because the other side would retaliate.

The informal cease-fires along the front indicates the ability of cooperation to arise when no altruism at all is involved. The generals on both sides who managed the war safely behind the lines deplored the refusal of troops to kill the enemy between the ordered offenses. They could and did issue orders telling their men that killing the enemy was to be a daily affair, but they could not be sure that these orders were in fact carried out. Only by accident did they discover a way to eliminate informal peace. The British began a policy of raiding the enemy, and a raiding party was expected to bring back prisoners or to suffer deaths trying. There was no way to fake these results, and as a result the cooperation that had arisen between enemies collapsed.

A final assumption we made in our story of the prisoners' dilemma was that only two individuals played. To see that the logic of the prisoner's dilemma is preserved and, indeed, strengthened when there are multiple players, consider the following classroom game. Your instructor divides your room into two groups. You must decide in each round of play whether to hold up a hand or a fist. Your score depends both on your decision and on how the majority of the other half of the class play. Table 4 shows the payoffs. Your goal is to score as many points as possible.

What will happen in this game? Perhaps you can try it and see. With most groups which play this game only a few hands are left after three or four rounds. (Notice that the problem of the commons, discussed in the last chapter, is identical to a prisoner's dilemma game with many people playing.)

TWO WAYS OF COORDINATION

The prisoner's dilemma indicates that if people act on the basis of self interest and there are no restrictions on their behavior, the results will not be in the interests of the group. The welfare of the group requires that some way be found to coordinate behavior and eliminate the conflicts involved in prisoner-dilemma situations. Groups have found two basic solutions.

One way involves establishing property rights and allowing markets to develop. Property rights make a decision maker consider all the costs and benefits of his decision. For example, consider the overgrazing problem in the last chapter. If each of the herders has to confine his cows to his own plot of land, then if he overgrazes the land, he will not be able to shift the costs of that decision to others who are using the land. As a result, he has an incentive to limit the size of his herd. Examining the way in which markets work will fill most of the the following chapters.

The second and more direct way is to have a strong central authority which regulates people's behavior and punishes deviations. Hunting regulations illustrate this method. Wild ducks are owned by no one. As a result, the problem of the commons should apply, and ducks should be hunted to extinction. They are not because governments set limits on when and where ducks may be hunted, and on how many ducks each hunter may kill. To the individual hunter these restrictions may seem as irritating limitations on his freedom. But without them, hunters as a group would be much worse off.

The use of central authority as a way of coordinating behavior is very widespread. It is the method which allows large business firms to exist. The boss coordinates his subordinates, directing them to actions which are in the interest of the all those who make up the firm. This solution is used in a more extreme form in the socialist economies of Eastern Europe and Asia. These economies are designed as if they were one giant firm. The bulk of economic decisions, such as what to produce and how to produce it, are made by government bureaucracies.

Which method, the market or central direction, is better? Evidence suggests that sometime the first is, and sometimes the second is. There are costs to using either method. As a result, no country totally relies on either. However, the Marxist nations have in the past struggled to eliminate coordination through property rights and the market. One of the important political changes of the 1980s has been the acknowledgement by the leadership of the Soviet Union and of China that their reliance on central coordination has been excessive.

Both ways of coordination rely on changing incentives, not people's goals. The basic idea in any set of incentives is that when some action is made more costly, people will do less of it. This simple idea has a multitude of complications when it is applied in the real world, which is fortunate for economists, or else they would not have nearly as much to study. In practice, virtually all attempts to change behavior by changing incentives produces changes more complex than those intended.

Those who designed the Soviet economic system began with a belief that "the problem with capitalism is that it produces for profit instead of for people's needs," and they set out to build a system that produces directly for people's needs and not at all for profit. "There was a period early in Soviet life when it was argued that Soviet worker and manager would work because of their enthusiasm for the revolution and their ideological fervor. That phase passed rather quickly."[3] Since use of markets violated Marxist ideology, there was only one system of coordination possible. A system of central planning evolved, a system in which all decisions about what people need are decided from the top.

---

[3] Marshall Goldman, *U.S.S.R in Crisis: The Failure of an Economic System*, (New York: W.W. Norton & Co., 1983).

To see how this system works, consider first how the operator of a shoe factory in the United States would make decisions. His major concern would be whether he could sell at a profit the shoes he made. In the Soviet Union, however, profit is of no concern to the manager of the state-owned shoe factory. Neither does he have to worry about selling the shoes. He only has to worry about producing what he is told to produce, and if he can do that, both he and the workers of the plant will receive sizable bonuses. The problem the Soviet Union has is that it is very difficult to specify in physical terms what the manager should do. (If you do not believe this, try to write down a set of instructions specifying what sort of shoes should be produced. Remember, instructions to produce "good shoes" or "attractive shoes" involve instructions that are not measurable.) The Soviet Union has produced huge numbers of shoes that no one will buy because they are of such low quality.

Or consider a nail factory. If it is told to produce as many nails as possible, it will only produce small nails. If told to produce as large a weight as possible, it will only produce very large nails. The Soviet Union has wasted billions of rubles searching for energy because it rewards drilling crews on the basis of the number of feet drilled. Because it is easier to drill many shallow wells than a few deep wells, drillers have drilled lots of shallow wells regardless of what was advisable geologically.

Incentives can be given whenever there are attempts to measure performance. In a series on the Soviet Union, the *Chicago Sun Times* reported a case in which a hospital had turned away a seriously ill patient because "they were nearing their yearly quota for patient deaths—and would be criticized by authorities if they exceeded it."[3] Certainly a good hospital should have fewer deaths than a bad hospital, other things constant. Thus it seems reasonable to ask hospitals for their death statistics, and these could be used to evaluate them. If they are evaluated on the basis of this statistic, they have the incentive to provide quality care. But they also have the incentive to avoid patients who are likely to die.

One need not go to the Soviet Union for examples of incentives with unintended results. An incentive problem occurred in the Job Corp program in the United States in the late 1960s which was much like the hospital example discussed in the previous paragraph. Job Corp was designed to provide job training for those who had very low levels of skills. In evaluating various centers that implemented the program, an attempt was made to see how many of those who went through the training got jobs. This statistic was used to determine how funding would be distributed in the future. On the surface this seems a very reasonable way to evaluate each training center. Those which are good should have had a higher placement rate for their graduates than those which were poor. However, this evaluation method had an undesirable consequence. Training centers began to try to attract people who were only moderately deficient in skills because they would be more likely to get a job and they began to discourage those who were severely deficient from the program.

A reason that designing effective incentives is so difficult is that information and knowledge are scarce. People have a limited capacity to know. In small groups people know a great deal about others simply from day-to-day interaction. But in large groups knowledge about others requires expenditure of time and effort. For a modern, complex economy to function well, people must coordinate their actions with the actions of many people, but all of these people have very limited knowledge of how their actions fit into the big picture. A good set of incentives gives people the essential information they need to know so that their decisions will mesh with others, and then encourages them to make those decisions. Not all sets of incentives, as we will see, do this task equally well, and none is without some problems.

---

[3] Nov 22, 1983, p. 4.

EXCHANGE AND POLITICS

The discussion of the prisoners' dilemma and the distinction between small groups and large groups suggests that economists are likely to have little to say about small-group situations. In fact, they have written very little, for example, about nomadic tribes or the interactions within the family. The realm of the economist has been in large-group situations in which individuals interact with others who they do not know well or at all. In these situations the assumption of self-interest seems to serve quite well even if it is not literally true. The area in which the analytical tools of economics have been most useful has been the area of exchange.

In small groups the role of exchange is replaced with gift-giving. A member of a family who has no food will be given food by those who do. But gift-giving usually has strings attached. People who receive gifts are expected to reciprocate, to give back something in the future. This reciprocity helps bind small groups together.

Exchange does not bind people together in the same way. Two people exchange only when both benefit. Neither incurs a social obligation as a result. In fact, where social obligations exist, exchange may not work well. Most people are uncomfortable negotiating a purchase from or sale to a close friend.

Exchange allows for extremely complex interactions among strangers. When you use a pencil, for example, you benefit from efforts of hundreds of people who in some way contributed to getting that pencil to you. Wood had to be grown, cut, and shaped. Graphite had to be mined, transported, and processed. Iron had to be mined, refined, and molded. The paint and eraser each required their own processes. All of the many people involved are probably total strangers to you.

Most of economics is devoted to discussing exchange, and the next chapter begins our discussion of it. Economists also spend time examining other topics. Of those other topics, the one which has drawn the most research is the study of government. However, the economic theory of government is less advanced than the theory of exchange for at least three reasons. First, often specific individuals are very important in government, and economics does not deal well with specific individuals. In monarchies and dictatorships, for example, the decisions of one individual may outweigh the decisions of all others. Economics seeks regularities in social life, and those regularities are more likely to occur when no one individual has appreciable effects on the group.

A second reason is that only fairly recently did economists realize that their analytical tools could be used to explain behavior in government. Economists usually assume that people are motivated by self-interest, but for many years they implicitly assumed that once the government employed a person, his motivations changed to unselfish and his knowledge became infinite. No economist ever stated this assumption; rather this was a hidden assumption in discussing the role of the government. This assumption allowed economists to treat the government as solution to problems. If, for example, the private market did not perform well, the government could remedy the problem. It was only in the 1960s and 1970s that economists fully realized that they had made what was for,them a most unusual assumption about behavior. Some of the most interesting and informative research undertaken in the 1970s was that which asked what would one expect if one assumes that those who work for the government are in fact no different from the rest of us, and that they seek only their own self-interest. This research suggested that the government might often make a bad situation worse.

A third reason is that the assumption of self-interest may not work as well in politics as in exchange. Politicians can and do make use of people's small-group responses. For example, many people feel loyalty to nation, party, or political personality. This loyalty can alter behavior from what it would be if based only on self-interest. We take a closer look at the topic of government and economic performance in Chapter 13

SUMMARY

This chapter considered the question of how individual motives and group results are related. It found that one cannot predict group results simply by knowing individual motives, but that group results also depend on incentives and the availability of information to those in the group. As a result, individual actions can have unintended group results. This chapter discussed two ways of coordinating decisions, and also explained why economists spend so much time discussing exchange, a topic to which we now turn.

QUESTIONS

1. The discussion in Chapter 2 suggested that there was a tendency to overgraze land open to all. (a) Suppose that there are only two herders but that each could control up to 300 cattle. Would this change the conclusion? (Hint: Use Table 3 from the the last chapter and draw a payoff matrix assuming that the strategies are "Graze 200" and "Graze 300." Then notice that this is a case in which there are many plays in the game, not just one.)

(b) Suppose that there are 100 herders as in the original example, but that the king demands 20% of milk production as a tax. What would he do if he wanted to maximize his tax receipts?

2. An arms race can be seen as a case of a prisoners' dilemma. Draw a payoff matrix showing the results of strategies "less armament" and "more armament" for two hostile nations. (Hint: Assume that war is a result of mistakes, that is, a nation does not go to war if they believe that they will lose. Payoffs then include "stalemate," "dominance," and "subservience.")

3. (A bit advanced) In the table below the numbers represent gains for Smith and losses for Jones. What is the best strategy for Smith? Why? What is the best strategy for Jones? Why? (Comment: This is a zero-sum game. Can you see why?)

Smith's Strategies

		A	B	C
Jones' Strategies	A	16	12	14
	B	2	10	9
	C	5	4	7

4. In 1800 buffalo were much more numerous in the United States than were cattle. By 1900 cattle were much more numerous than buffalo. What is the key to explaining the different fates of the buffalo and cattle? (Comment: The answer which this chapter suggests is incomplete. The U.S. Army also encouraged the destruction of buffalo in its attempt to defeat the Plains Indians.)

5. Some years ago the farmers in one of the Caribbean nations were having a problem with snakes which ate chicken eggs. To solve the problem, the government instituted a bounty on the snakes. The program worked well at first, but then the authorities noticed that rather than decreasing, the number of snakes being turned in was increasing. Investigating, they found that the bounty had been set too high. What unwanted incentive would very high bounties give people?

6. (a) Using the definitions from the last chapter, how would you classify each of the four possible outcomes in Table 3: as stable equilibrium, as unstable equilibrium, or as a situation of disequilibrium? Explain your reasons for each choice.

(b) The definition of equilibrium which economists use in game theory is Nash equilibrium. A Nash equilibrium (named after the economist who first defined it) exists when each player finds that there is no way to improve his results given the strategies played by the other player(s). Where will the Nash equilibrium be in Table 3?

7. Primitive societies of hunter-gatherers have neither property rights nor strong central authorities. (a) How can they avoid the problems of coordination and conflict over scarce resources? (Hint: They also have very low population densities. What does this do to a problem of the commons?)

(b) Does your answer suggest a theory about how and when property rights come into being?

8. Evaluate the following statement: "The problem with capitalism is that it produces for profits instead of for people's needs."

9. Can you switch two numbers in Table 4 so that self-interest is harmonized with group interest?

CHAPTER 4

THE MODEL OF SUPPLY AND DEMAND

After you read this chapter and work the problems at the end, you should be able to:

Define demand, supply, inferior good, normal good, substitute, complement, law of demand, price taker, price searcher, market-clearing price.

Distinguish between changes in demand and changes in quantity demanded.

Distinguish between changes in supply and changes in quantity supplied.

Predict how changes in factors such as income, prices of substitutes, prices of inputs, etc. affect the supply and demand curves and equilibrium quantity and price.

Explain why we can treat the demand curve as positions of buyer equilibrium and the supply curve as positions of seller equilibrium.

The last three chapters have told you that economists analyze the results of people's actions when these actions are based on self-interest. When actions are based on self-interest, people respond to incentives, that is, to costs and benefits. When the costs of an activity are raised, or the benefits reduced, people do less of the activity. In examining what happens in a market, economists have found that they can use this simple idea of action based on costs and benefits to construct a model (or theory) which can explain how many markets work. This model is the model of supply and demand. It is perhaps the most basic of the models economists use to explain the world around them.

Given the model's importance in the way modern economists think, it is surprising that one does not find the model in the writings of Adam Smith, David Ricardo, Thomas Malthus, or John Stuart Mill, though all of these pioneers in economics used the words "supply" and "demand" frequently. The modern supply-and-demand model does not appear until 1890, when Alfred Marshall published his *Principles of Economics*.

DEMAND

What determines the amount of a product that people are willing to buy during some period of time? For example, what determines the amount of hamburger purchased in Chicago during a week? Economists answer such questions by examining the costs and benefits of buying the product. When any of the costs or benefits change, the amount of the product which people will buy should also change.

The benefits a person gets from a product depend on his goals. These goals are referred to in many ways in discussions of demand. The words "tastes," "wants," "needs," "preferences," and "usefulness" all refer to goals. When peoples' goals change, the amount of benefit they get from the good changes, and this will cause them to change the amount of the good they want to buy.

Goals (or preferences or tastes) depend on many factors, such as the age of people and the amount of education they have had. Social custom is an important determinant of preferences and can account for many differences in demand among groups. One can explain the large differences in squid sales in Japan and the United States, or the large differences in consumption of horse meat in Europe and the United States, almost entirely in terms of differences in preferences caused by differences in social custom.

The most obvious cost a person bears in buying a product is the price of the product. Price reflects cost because people have a limited amount of funds which they can spend, and if they spend their money on one thing, they cannot spend it on another. When the price of a product goes up, the amount of other things which a person must give up in order to buy the product rises. As a result, we expect people to buy more hamburger if the price is $1.00 per pound than if it is $2.00 per pound.

It follows that the amount of income people receive should also be important. The amount of income determines the options which a person must give up when he buys a product. If a person with a low income spends $5000 for a trip around the world, he will have to cut back on food, clothing, or shelter. The same trip will cause a person with a high income to cut back on a very different set of options.

Increases in people's incomes lead to more consumption of most products. These products are called **normal** goods. There are some products, however, which people use less of as their income increases, and these products are called **inferior** goods. Public transportation is an example—as people's incomes rise, they stop riding the bus and drive their own cars. Blue jeans were once another example—people with higher incomes bought them less frequently than people with lower incomes. It was because they were a symbol of "working class" clothes that they were adopted by the radical left in the 1960s, and from there they moved into high fashion.

Prices of related goods are also important.

Goods which are **substitutes** satisfy the same set of goals or preferences. An example of a substitute for hamburger is pork. If pork prices are high, people are tempted to shift away from pork to hamburger, and if pork prices are low, people are tempted to shift from hamburger to pork. The opposite of a substitute is a **complement**, a good which helps complete another in some way. Catsup and hamburger buns are complements to hamburger, and if they are priced low enough, consumption of hamburger may rise. Sometimes goods are such good complements that they are sold together and we think of them as a single item. Left shoes and right shoes are an example.

There are other factors that can influence the amount of a particular product that people are willing to buy, such as the number of consumers in the market and their expectations about future prices, incomes, and quality changes. To get a complete list for any product might be time consuming and difficult, but it is not necessary because we want to consider only the relationship between price and the quantity of a product that people are willing to buy during some interval of time. To do this we will assume all other factors are held constant.

TABLE 1
A DEMAND CURVE

PRICE OF WIDGETS	NUMBER OF WIDGETS BOUGHT
$1.00	100
$2.00	90
$3.00	70
$4.00	40

The relationship between price and the amount of a product people want to buy is what economists call the **demand curve**. This relationship is inverse or indirect. This means that as price gets higher, people want less of a particular product. This inverse relationship is almost always found in studies of particular products, and its very widespread occurrence has given it a special name: **the law of demand**. The word "law" in this case does not refer to a bill that the government has passed but to an observed regularity.

There are various ways to express the relationship between price and the quantity that people will buy. Mathematically one can say that quantity demanded is a function of price, with other factors held constant, or:

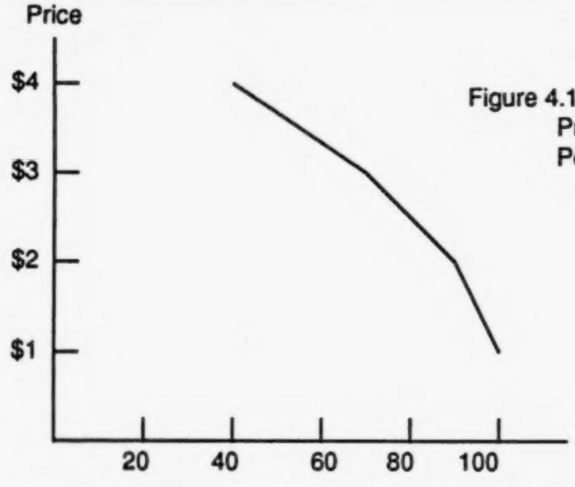


Figure 4.1 A Demand Curve Shows How Price Influences the Amount People Want to Buy

TABLE 2
A DEMAND CURVE CAN SHIFT

PRICE	QUANTITY
$1.00	[100] becomes 80
$2.00	[90] becomes 70
$3.00	[70] becomes 50
$4.00	[40] becomes 10

Qd = f(Price, other factors held constant)

A more elementary way to capture the relationship is in the form of a table. The numbers in Table 1 are what one expects in a demand curve: as price goes up, the amount people are willing to buy decreases. (A widget is an imaginary product that some economist invented when he could not think of a real product to use in illustrating an idea.) The same information can also be plotted on a graph, where it will look like the graph in Figure 1. (To follow mathematical convention, quantity should be on the vertical axis since quantity depends on price. But Alfred Marshall put price on the vertical axis when he developed the modern treatment of supply and demand in 1890, and his way has become tradition.)

If one of the factors which is being held constant becomes unstuck, changes, and then is held constant again, the relationship between price and quantity will change. For example, suppose the price of getwids, a substitute for widgets, falls. Then people who previously were buying widgets will reconsider their choices, and some may decide to switch to getwids. This would be true at all possible prices for widgets. These changes in the way people will behave at each price will change the demand curve to look like Table 2 and Figure 2.

If the price of widgets is originally $1.00 and people are buying 100, they may change to 90 for two reasons. One reason is that the price may rise to $2.00. The other reason is that one of the factors which is assumed constant may change, so that even though the price has not changed, quantity will. Economists distinguish these two cases. In the first case the demand relationship or schedule has not changed, but there has been movement within the relationship. Economists call a change of this sort a **change in quantity demanded.** The second sort of change is an alteration of the relationship. The original pairing of price and quantity is destroyed and replaced by a new pairing.

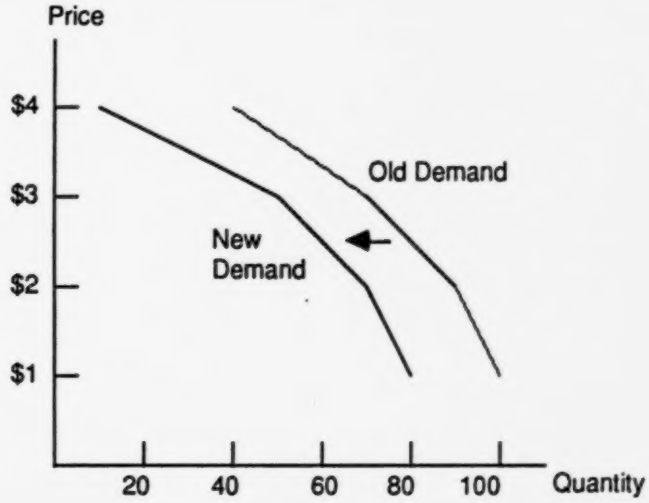


Figure 4.2 A Shift In Demand

Economists call this sort of change a **change in demand**.

It is important to realize that though the demand relationship looks concrete when it is illustrated with a table or graph, in everyday life demand curves are hidden. A demand curve refers to what people would do if various prices were charged, and very rarely are enough prices charged so a clear demand curve can be seen. This is not to say that the concept is of no importance to people who sell. They may not be interested in the demand curve as a relationship, but they do find it a boundary or constraint on their behavior. If there were an actual widget seller facing the demand curve in Table 1, he would find that he could not sell more than 90 widgets if he wanted to charge $2.00. He could of course sell fewer if he wanted to. He could sell only 70 at $2.00, but if he did this he would earn far less than he could. If he wanted to sell more than 90, he would have to lower his price.

Thus, to an actual businessman the demand curve is important as a limitation on what he can do. A businessman may not know exactly where the demand curve is, and he may not think of it as fixed. Advertising—either informing or persuading people—can move the boundary. As we proceed further, we will see that there are still other ways to view the demand curve in addition to seeing it as a mathematical relationship and as a boundary limiting sellers.

SUPPLY

What determines the amount of a good or service that people are willing to sell during some period of time? We saw in Chapter 2 that people sell things because it is a way, indirect but effective, of obtaining other things which they prefer. Sellers intend to make a profit from their sales, and economists assume that they want their profits as large as possible. Since profit is the difference between benefits in the form of revenues and costs, anything which influences revenues or costs can influence the amounts sellers want to sell.

Revenue is found by multiplying the price of the product by the amount sold. A change in price changes revenues, and hence profits, so it is a major determinant of the amount sellers will want to sell. Since a higher price leads to higher profit, and a higher profit leads to a larger amount that sellers will want to sell, one expects that a greater quantity should be supplied when price is higher. Thus, the relationship between quantity that sellers will sell and price should be direct or positive. Though this is almost always the case, there are a few exceptions. An example is labor; as wages go up, people may decide to enjoy their higher wages and work less. As a result, there is no law of supply that matches the law of demand.

The cost of something is what must be given up in order to get it. Often these costs are only monetary, in which case they are easy to see. If the price of an input increases, the cost of the output will increase, and, other things held constant, profits will decrease. The seller will then have to decide if shifting part of his resources and effort to other products will improve his well-being.

Production costs are determined not only by the prices of inputs, but also by **technology**. Technology represents the knowledge of how inputs (such as labor, raw materials, energy, and machinery) can be combined to produce the product. If this knowledge increases, so that people find cheaper ways to make the same output, then, other things held constant, profit would increase and we expect sellers to respond by producing more.

Costs may be nonmonetary as well as monetary. For example, a farmer who plants corn takes the expected price of soybeans into account in deciding how much corn to plant. If soybeans are expected to sell for a high price, then the farmer may find that shifting some of his land from corn production to soybean production will increase profit. The decision to plant corn means that the farmer gives up the opportunity to plant soybeans (as well as giving up the money for seed, fuel, equipment, and labor). Since we have defined cost as what must be given up

TABLE 3 A SUPPLY CURVE

PRICE	QUANTITY SUPPLIED
$1.00	10
$2.00	40
$3.00	70
$4.00	140

to get something, the prices of other goods that sellers could otherwise produce and sell must be part of the calculation of the cost of production.

There are other factors that can influence the amount of a product that sellers will sell, such as the number of sellers, expectations about the future, and whether or not there are by-products in production that are valuable. (An example of a valuable by-product is cottonseed in the production of cotton. A farmer who produces cotton also gets cottonseed, which yields cottonseed oil, a widely-used vegetable oil.) But as in the discussion of demand, the emphasis in the discussion of supply is on the relationship between quantity and price. To focus on this relationship, all other factors must be assumed to be constant.

There is another assumption that has not yet been clearly stated though it applies to demand as well as to supply. For a supply curve to exist, there must be a large number of sellers in the market, and for a demand curve to exist, there must be many buyers.. In both cases there must be enough so that no one believes that what he does will influence price. In terms which were first introduced into economics in the 1950s and which have become quite popular, everyone must be a **price taker** and no one can be a **price searcher**. A price searcher is one with ability to set the price in a transaction, while a price taker cannot influence the price but must take or leave it. The ordinary consumer knows the role of price taker well. When he goes to the store, he can buy one or twenty gallons of milk and it does not change the price. The assumption that both buyers and sellers are price takers is a crucial assumption, and often it is not true with regard to sellers. When it is not true, caution must be taken when using the supply and demand model which lies in the pages ahead.

The relationship between the quantity sellers want to sell during some time period (quantity supplied) and price can be expressed mathematically in functional form as:

Qs = f(price, other factors held constant).

It can also be illustrated in the form of a table or a graph, as in Table 3 and Figure 3. The graph has a positive slope, which is the slope one normally expects from a supply curve.

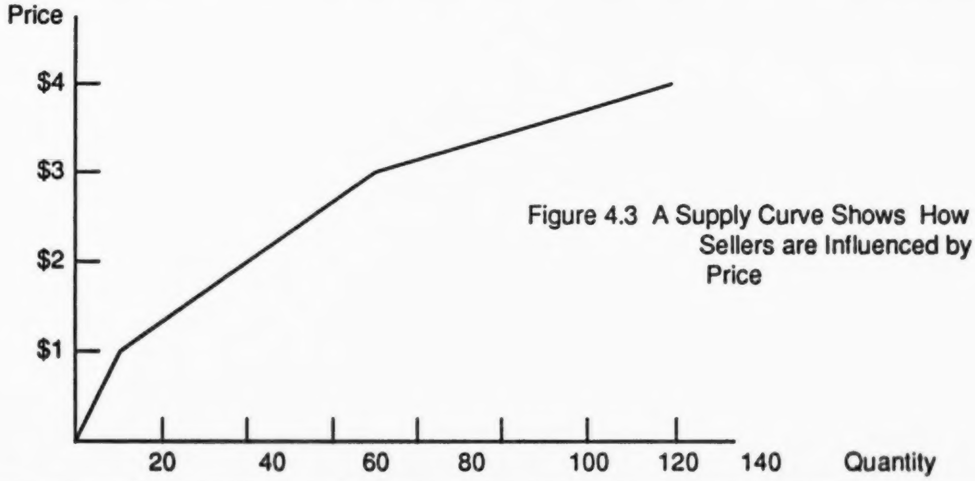


Figure 4.3 A Supply Curve Shows How Sellers are Influenced by Price

TABLE 4
A SUPPLY CURVE CAN SHIFT TOO

PRICE	QUANTITY SUPPLIED
$1.00	[10] becomes 20
$2.00	[40] becomes 60
$3.00	[70] becomes 100
$4.00	[140] becomes 180

If one of the factors which is held constant changes, the relationship between price and quantity, which is supply, will change. If the price of an input falls, for example, the supply relationship may change as in Table 4 and Figure 4.

As with demand, economists separate changes in the amount that sellers will sell into two categories. A **change in supply** refers to a change in behavior of sellers caused because a factor held constant has changed. As a result of a change in supply, there is a new relationship between price and quantity. At each price there will be a new quantity and at each quantity there will be a new price. A **change in quantity supplied** refers to a change in behavior of sellers caused because price has changed. In this case the relationship between price and quantity remains unchanged, but a new pair in the list of all possible pairs of price and quantity has been realized.

Supply curves as well as demand curves appear much more concrete on an economist's graph than they appear in real markets. A supply curve is mostly potential—what will happen if certain prices are charged, most of which will never be charged. From the buyer's perspective, the supply curve has more meaning as a boundary than as a relationship. The supply curve says that only certain price-quantity pairs will be available to buyers, those lying left of the supply curve.

THE MODEL OF SUPPLY AND DEMAND

To this point we have developed two behavioral statements, or assertions about how people will act. The first says that the amount buyers will buy depends on price and other factors which are assumed constant. The second says that the amount sellers will sell depends on price and other factors which are assumed constant. In mathematical terms our model is:

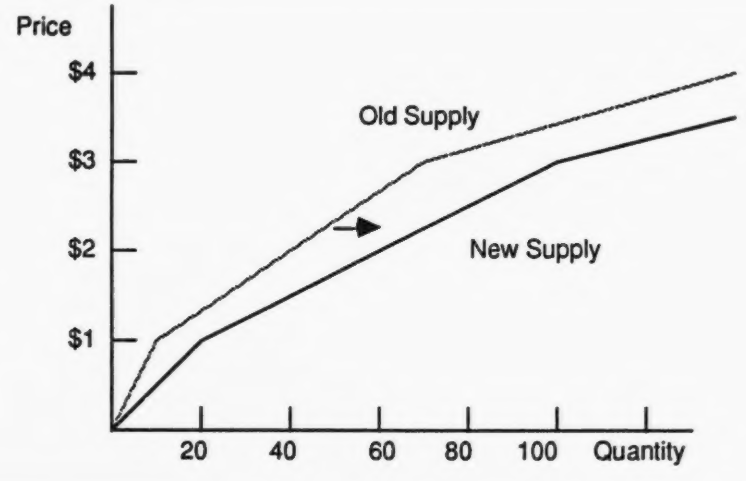


Figure 4.4 A Shift in Supply

TABLE 5
SUPPLY AND DEMAND TOGETHER AT LAST

PRICE	Qd	Qs
$1.00	100	10
$2.00	90	40
$3.00	70	70
$4.00	40	140

Qd = f(price, constants)

Qs = g(price, constants).

This is not a complete model. Mathematically the problem is that we have three variables (Qd, Qs, price) and only two equations, and this system will not have a solution. To complete the system we need to add another equation, and this will be an equilibrium condition. The equation is very simple:

Qd = Qs.

It says that equilibrium exists if the amount sellers are willing to sell is equal to the amount buyers are willing to buy.

Consider Tables 1 and 3 and Figures 1 and 3. If we combine them, we get the Table 5 and Figure 5. It should be obvious that the price of $3.00 is the equilibrium price and the quantity of 70 is the equilibrium quantity. At any other price, sellers would want to sell a different amount than buyers want to buy. On the graph, the equilibrium price and quantity are indicated by the intersection of the supply and demand curves.

If one of the many factors which is being held constant changes, then equilibrium price and quantity will change. Further, if we know which factor changes, we can often predict the direction of changes, though rarely the exact magnitude. For example, the market for wheat fits the requirements of the supply and demand model quite well. Suppose there is a drought in the main wheat producing areas of the United States. How will we show this on a supply and demand graph such as Figure 5? Should we move the demand curve, or the supply curve, or both? What will happen to equilibrium price and quantity?

A dangerous way to answer these questions is to first try to decide what will happen to price and quantity and then decide what will happen to the supply and demand curves. This is a route to disaster. Rather one must first decide how the curves will shift, and from the shifts in the curves decide how price and quantity would change.

What should happen as the result of the

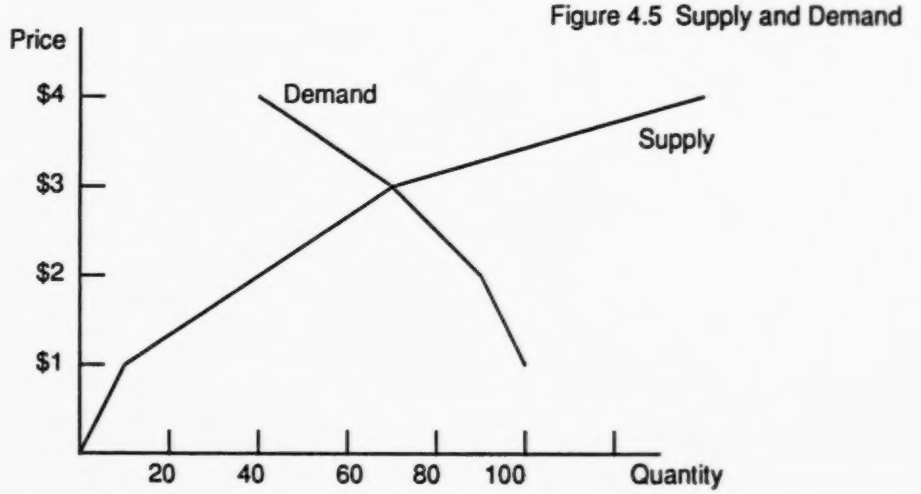


Figure 4.5 Supply and Demand

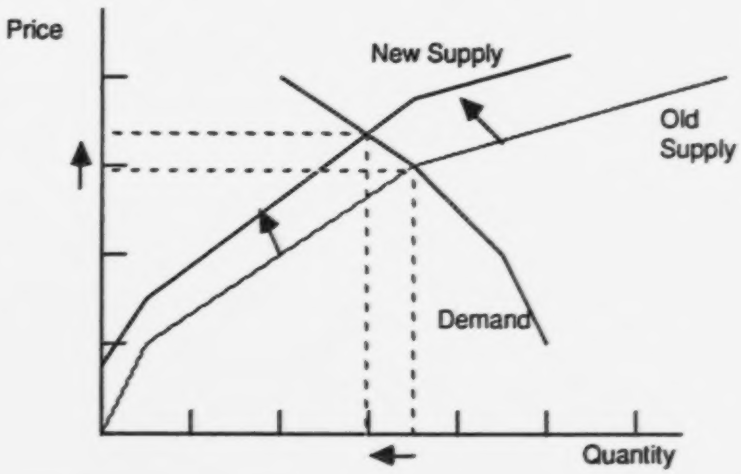
drought? One begins by asking whether buyers would change the amount they purchased if price did not change and whether sellers would change the amount sold if price did not change. On reflection one realizes that this event will change seller behavior at the given price, but is highly unlikely to change buyer behavior (unless one assumes that more than the drought occurs, such as a change in expectations caused by the drought). Further, at any price, the drought will reduce the amount sellers will sell. Thus, the supply curve will shift to the left and the demand curve will not change. There will be a change in supply and a change in quantity demanded. The new equilibrium will have a higher price and a lower quantity. These changes are shown in Figure 6.

What should one predict if a new diet calling for the consumption of two loaves of whole wheat bread sweeps through the U.S.? Again one must ask whether the behavior of buyers or sellers will change if price does not change. Reflection should tell you that it will be behavior of buyers which would change. Buyers would want more wheat at each possible price. The demand curve shifts to the right, which results in higher equilibrium price and quantity. Sellers would also change their behavior, but only because price changed.

The model of supply and demand requires that buyers and sellers be clearly defined groups. Notice that in the list of factors which affected buyers and sellers, the only common factor was price. Few people who buy hamburger know or care about the price of cattle feed or the details of cattle breeding. Cattle raisers do not care what the income of the buyers is or what the prices of related goods are unless they affect the price of cattle. Thus, when one factor changes, it affects only one curve, not both. When buyers and sellers cannot be clearly distinguished, as in the New York Stock Exchange, where the people who are buyers one minute may be sellers the next, one cannot talk about distinct and separate supply and demand curves.

The model of supply and demand also assumes that both buyers and sellers have good information about the products qualities and availability. If information is not good, the same product may sell for a variety of prices. Often, however, what seems to be the same product at different prices can be considered a variety of products. A pound of

Figure 4.6 Shifting the Supply Curve Moves Us along the Demand Curve



hamburger for which one has to wait 15 minutes in a check-out line can be considered a different product than identical meat that one can buy without waiting. Chapters 12 and 16 look at cases in which information is not good.

In addition, the supply and demand model needs well-defined property rights. Chapter 3 suggests this conclusion in its discussion of how property rights and markets provide one way of coordinating decisions. When property rights are not clearly defined, the seller may be able to ignore some of the costs of production, which will then be imposed on others. Alternatively, buyers may not get all the benefits from purchasing a product. Others may get some of the benefits without payment. Chapters 12 and 14 will consider these cases.

Finally, as was noted above, the supply and demand model requires many buyers and sellers. If there is only one seller, the seller can search along the demand curve of the buyers for the position that is most profitable. In this case it is not just price that matters, but the slope of the demand curve as well. The seller in this case is not a price taker, but a price searcher. Chapter 15 will examine how markets with price searchers function.

Even if the assumptions underlying supply and demand are not met exactly, and they rarely are, the model often provides a fairly good approximation of a situation, good enough so that predictions based on the model are in the right direction. This ability of the model to predict even when some assumptions are not quite satisfied is a reason economists like the model so much.

THE ADJUSTMENT PROCESS

We have developed the model of supply and demand as an equilibrium model. We have said nothing about how adjustments from disequilibrium to equilibrium take place. To develop this idea, it is useful to take still another view of supply and demand curves, to view demand as points of buyer equilibrium and supply as points of seller equilibrium.

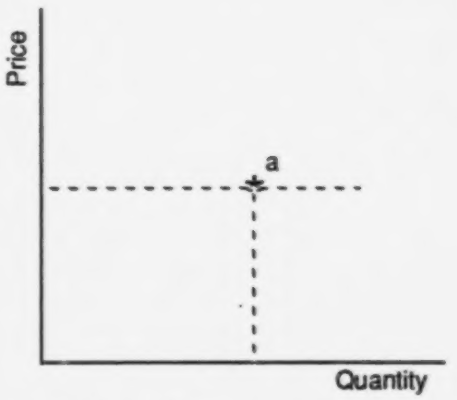
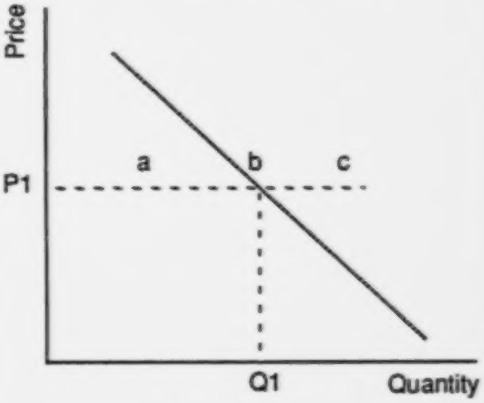


Figure 4.7 Is Point *a* on the Demand Curve?

Suppose price is at P1 in Figure 7. Will point a be a point on the demand curve? If people would like to buy more than Q1 at price P1, point a must lie to the left of the demand curve. In this case some consumers are unhappy with the amount they have purchased, and will try to purchase more. If there is no more to purchase, some will attempt to offer more money for the product, or they will increase the time they devote to getting the product. The important idea is that if point a lies to the left of the demand curve, people will be unhappy with their situation and will change their behavior. If point a lies to the right of the demand curve, people will decide that they are buying too much of the product and will cut purchases. Again, if point a is not on the demand curve, people will change their behavior.

This analysis indicates that positions off the demand curve are not positions of buyer equilibrium. They also help explain why the demand curve can be considered a boundary. Though buyers may not be in equilibrium at point a in Figure 8, they can be held there and they can adjust in ways that do not change the money price. They cannot be held

Fibure 4.8 The Demand Curve Shows Buyer Equilibirum



at point b unless there is some way to force people to buy a product when they do not want it.

Point c in Figure 8 is a position of buyer equilibrium because given price P1, people will be satisfied with Q1 and will do nothing to change their behavior. Buyers would, of course, prefer a lower price than P1—they are always willing to move down the demand curve. However, this is not the issue here. Given P1, Q1 is the preferred quantity.

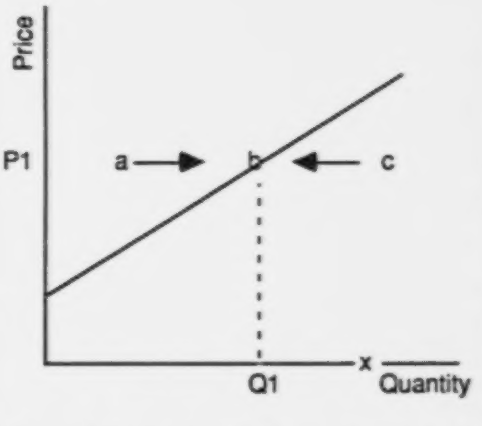


Figure 4.9 The Supply Curve Shows Seller Equilibrium

A similar analysis holds for the supply curve. If at point a in Figure 7 suppliers find that they could increase profits (or reduce losses) by moving to the right from Q1 to a larger quantity, they are not in equilibrium and they are not on the supply curve but to the left of it. If they find that they could increase profits by cutting production, they are to the right of the supply curve and out of equilibrium. As Figure 9 shows, if price is fixed at P1, sellers will try to adjust toward the point on the supply curve.

It is possible to force a seller to a position to the left of the supply curve. This is the case in which the seller would like to sell more at the given price, but for some reason can not. One reason might be that the buyers will not buy as much as the sellers would like to sell. It is virtually impossible—short of slavery—to force sellers to the right of the supply curve. This is the situation in which sellers are selling more than they want to at the given price. Thus, the supply curve represents a boundary facing the buyers. If buyers could force sellers to the right of the supply curve, they would find it advantageous to force sellers to a position such as x in Figure 9, which represents getting something for nothing.

Sellers prefer higher prices to lower prices. Though all points on the supply curve represent points of equilibrium, not all are equally preferred by sellers.

The above analysis helps explain how an adjustment process takes place in the supply and demand model. If price is originally P1 in Figure 10, only Q1 will be sold even though buyers would like to buy Q2. The difference Q2 - Q1 represents a shortage. The sellers are in equilibrium in this situation, but buyers are not. Some buyers who cannot obtain the product are willing to offer more, and sellers are always willing to accept a higher price. Therefore, the actions of the buyers, as they compete with each other to obtain the amount that is available, drive the price upward in this model toward market equilibrium.

If price is originally at P2 in Figure 11, only Q1 will be sold because this is all that buyers

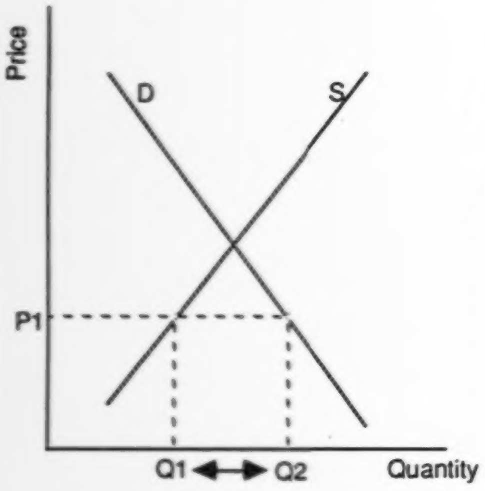


Figure 4.10 A Shortage

will purchase, even though sellers are willing to sell more. The difference Q2 - Q1 is called a surplus. In this situation the buyers are in equilibrium but the sellers are not. Sellers will compete among themselves to get rid of the surplus, and some will be willing to offer their product at a lower price. Buyers are always willing to move down the demand curve, so there is a tendency to move downward toward market equilibrium in Figure 11.

One would expect that if left to itself a market that can be represented by supply and demand would tend to adjust to the point where the supply and demand curves cross. The price at this intersection is called the market-clearing price. There is, however, the possibility that the existence of lags in the adjustment process may make the adjustment more complex than the previous discussion indicates.

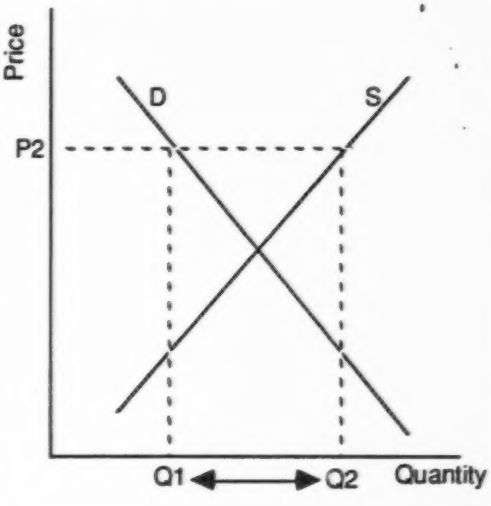
Suppose that the price of cattle feed rises sharply. This is a factor which should affect the supply curve of cattle and shift it to the left. The profitability of cattle production is reduced at each possible price, and some producers will drop out of the industry while others will curtail production. Looking at the curves, we see that prices should rise and quantity drop.

But initially prices might drop and quantity rise, which is exactly opposite the prediction from the supply and demand graph. The higher costs of feed will encourage farmers to raise fewer cattle, but as part of that cutback, they will temporarily send more cattle to the slaughter houses. The prediction which supply-demand analysis gives will ultimately be correct, but it will not be correct in the process of adjustment.

More complicated adjustment patterns are possible. Suppose, for example, that higher beef prices shift the demand for pork to the right. Supply and demand analysis says that this should increase pork prices, and at the higher prices, farmers should produce more hogs.

However, hog production takes time, and will only happen if farmers expect the higher prices to continue for a long time. If pork producers do expect the higher prices to last, they may decrease the number of pigs sent to slaughter, further increasing price. A sow can either produce pork or baby pigs, but not both. If farmers expect high prices to last,

Figure 4.11 A Surplus



they will keep their sows for piglet production.

In six months to a year, the baby pigs will have grown enough to go to market. If enough farmers had expected the high prices to last, they may have produced so many pigs that pork prices will now plunge to a level below that which is considered normal. The new, abnormally low price can then influence decisions which will not affect the price for many months. You should see that, once disturbed, a market with long time lags in production may bounce around for years before it finally finds its way back to equilibrium. If such a market is disturbed often enough, its prices and quantities will never come to rest at equilibrium levels.

Microeconomics assumes that markets clear, that is, they are always in equilibrium. Its analysis begins with the assumption that equilibrium has been reached, and then asks questions about that equilibrium. However, adjustment problems are very important in macroeconomics. Macroeconomics cannot assume there are no adjustment problems, or else it assumes away one of the problems it wants to explain, unemployment. In fact, much of macroeconomics is about the forces which bump an economy away from equilibrium, and why, once it is away, it has problems in reaching a new equilibrium.

SUMMARY

This chapter introduced you to one of the most common models which economists use, the model of supply and demand. It showed that buyers, basing their behavior on the weighing of costs and benefits, will buy more as the price of a product decreases, and that sellers, also basing their behavior on the weighing of costs and benefits, will usually be expected to sell less as the price of a product decreases. The interaction of buyers and sellers in a market establishes an equilibrium price and quantity which can be illustrated graphically with supply and demand curves.

QUESTIONS

1. A profit-maximizing professor asks her class the maximum they would pay for an A. These are the numbers she gets: $100, $50, $75, $0, $200, $0, $500, $50, $50, $10, $25, $75, $50. Construct a table which shows the demand curve for As in this class.

2. Assume that normal supply and demand curves can represent the American market for beef. As a result of each event below, indicate in which direction the supply and demand curves will move (right or left) and what happens to equilibrium price and quantity.

	SUPPLY	*DEMAND*	*Q*	*P*
A. The price of cattle feed rises.	___	___	__	__
B. People start worrying about the sanitation in meat slaughtering houses.	___	___	__	__
C. A severe winter destroys 5% of the herds.	___	___	__	__
D. The ban on the use of DES (a growth hormone used for cattle) is lifted.	___	___	__	__

E. In an effort to aid farmers, government grants a subsidy of $.05 a pound for beef produced. —— —— — —

F. Government grants farmers a subsidy of $10 per acre of land they own. —— —— — —

G. All foreign beef is excluded from the American market. —— —— — —

H. Per capita income rises. —— —— — —

I. Housewives expect rising beef prices. —— —— — —

J. Food companies improve the taste of their soybean meat substitutes. —— —— — —

K. The price of ham falls. —— —— — —

L. A cheap and perfect substitute for leather is developed. —— —— — —

M. The price of milk decreases and dairy farmers sell their herds. —— —— — —

N. The price of ketchup (a complementary good) rises. —— —— — —

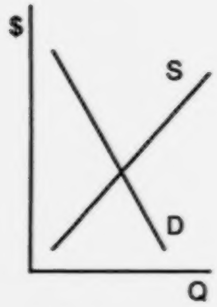
Note: for the last three events, determine the price and quantity which will appear in the market. They may not be equilibrium prices or quantities.

O. The government establishes a price ceiling above equilibrium price. —— —— — —

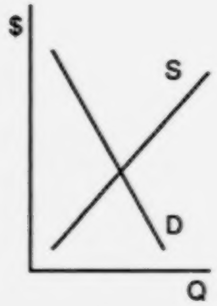
P. The government establishes a price ceiling below equilibrium price. —— —— — —

Q. The government sets a price floor above equilibrium price. —— —— — —

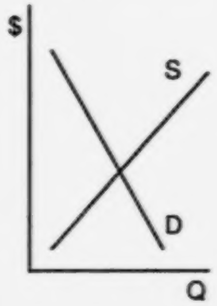
Warning: Several of the items above are slightly ambiguous and the answer you give depends on the assumptions you make.



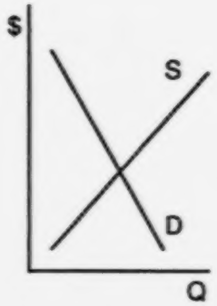
Shirts made of Synthetic Fiber



Shirts made from Cotton



Cotton

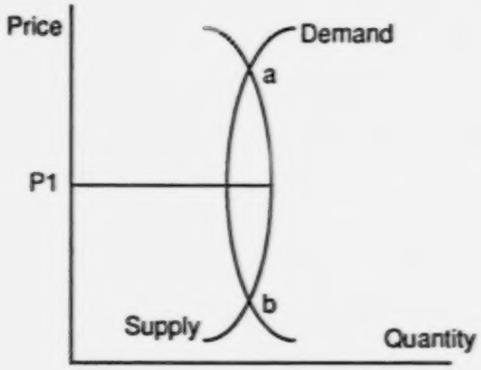


Cotton-seed Oil

3. Suppose the markets shown above may be described by supply and demand curves. Show what should happen in each market as a result of a sharp increase in the price of petrochemicals from which synthetic fibers are made. Add a note of explanation to each graph. (Hint: work from left to right and ignore any feedback.)

4. "If the price goes down, the seller will have to sell more to make the same profit. Hence the supply curve should have a negative slope just as the demand curve does." What is wrong with this reasoning? (Hint: What did we assume about the level of profits a firm seeks? According to this quotation, what would happen to profits if a seller sells more at the original, high price?)

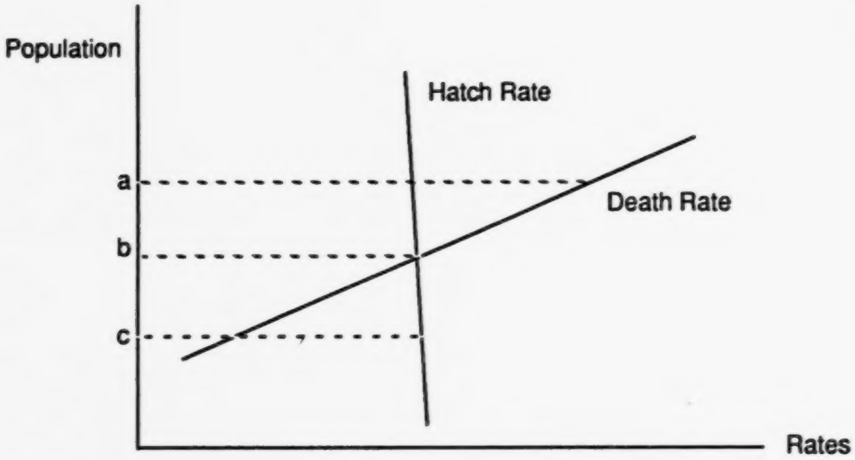
5. Suppose price is at P1 in the figure below. What will the adjustment process cause to happen? Is there any difference between equilibria a and b?



6. Suppose an ecologist is studying the population of the golden-throated song sparrow and he finds the relationships illustrated in the graph below between the hatch and death rates and population, given that existing food supply, level of predation, etc., are fixed. a) What will happen if the population is at level a? At level b? At level c?

b) Show on the graph what will happen if the food supply increases.

c) Explain what will happen on the graph if, due to an increase in the use of pesticides, fewer eggs hatch.



d) Explain what will happen on the graph if the number of predators increases.

e) In what ways do the mechanics of this model resemble the mechanics of the supply and demand model?

7. The law of demand says that as something gets cheaper, people buy more of it and as something gets more expensive, people buy less. Which of the following events would increase the amount of college education people buy?

a. tuition rates fall.
b. full-time jobs available to high school graduates decrease.
c. the price of food falls.
d. the price of housing falls.

Costs are whatever must be given up to get something. Which of the following are costs of going to college: tuition and fees, book expenses, entertainment expenses, housing costs, food costs, lost chances of earning income? How much does it cost you to go to college using this definition of cost?

CHAPTER 5

ELASTICITY AND REVENUE

After you read this chapter and work the problems at the end, you should be able to:

Compute price elasticities of supply and demand when given the curves in the form of a table.

Explain what is meant when one says demand is elastic or when one says demand is inelastic.

Define income and cross-price elasticity, and explain what they measure.

Compute marginal revenue when given total revenue, and vise versa.

Compute average revenue when given total revenue, and vise versa.

Explain why marginal revenue is the slope of the total revenue curve.

Recognize the area that represents total revenue on a demand or supply graph.

This chapter explores some economic terms and concepts which follow directly from supply and demand curves and which are important building blocks for later chapters. It begins with the concept of elasticity, which measures how people respond to changes. It goes on to concepts of average, total, and marginal revenue, and finishes by linking revenue and elasticity.

PRICE ELASTICITY

Businesses know that they face demand curves, but rarely do they know what these curves look like. Yet sometimes a business needs to have a good idea of what part of a demand curve looks like if it is to make good decisions. If Rick's Pizza raises its prices by ten percent, what will happen to its revenues? The answer depends on how consumers will respond. Will they cut back purchases a little or a lot? This question of how responsive consumers are to price changes involves the economic concept of **elasticity**.

Elasticity is a measure of responsiveness. Two words are important here. The word "measure" means that elasticity results are reported as numbers, or elasticity coefficients. The word "responsiveness" means that there is a stimulus-reaction involved. Some change or stimulus causes people to react by changing their behavior, and elasticity measures the extent to which people react.

The most common elasticity measurement is that of **price elasticity of demand**. It measures how much consumers respond in their buying decisions to a change in price. The basic formula used to determine price elasticity is:

$$(1)\ e = \frac{\text{percentage change in quantity}}{\text{percentage change in price.}}$$

If price increases by 10%, and consumers respond by decreasing purchases by 20%, then equation one says that the elasticity coefficient is -2. The negative sign is there because an increase in price (a positive number) leads to a decrease in purchases (a negative number). However many economists ignore the negative sign, as we will in the following discussion.

An elasticity coefficient of 2 shows that consumers respond a great deal to a change in price. If, on the other hand, a 10% change in price causes only a 5% change in sales, the elasticity coefficient will be only 1/2. Economists would say in this case that demand is **inelastic**. Demand is inelastic whenever the elasticity coefficient is less than one. When it is greater than one, economists say that demand is **elastic**.

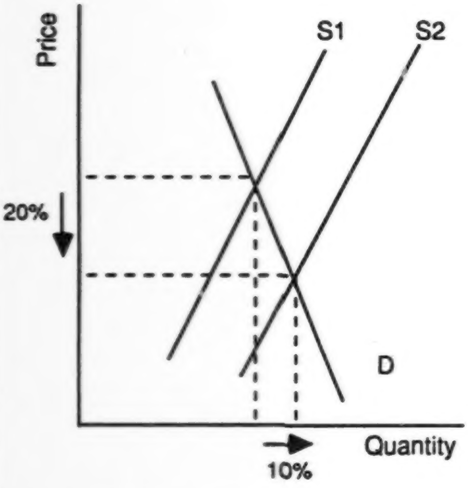
Goods which have few good substitutes generally have a lower elasticity of demand than goods with many substitutes. This also means that more broadly defined products have a lower elasticity than narrowly defined products. The price elasticity of demand for meat will be lower than the price elasticity of pork, and the price elasticity for soft drinks will be less elastic than the price elasticity for colas, which in turn will be less elastic than the price elasticity for Pepsi.

Time plays an important role in determining both consumer and producer responsiveness for many items. The longer people have to make adjustments, the more adjustments they will make. When the price of gasoline rose rapidly in the late 1970s as a result of the OPEC cartel, the only adjustment consumers could initially make was to drive less. With time they could also move closer to work or find jobs closer to home, and switch to more fuel-efficient cars.

Just as we can measure how responsive buyers are to a change in price, we can measure how responsive sellers are. This measurement is the **price elasticity of supply.** It has the same formula as price elasticity of demand, only the quantity in the formula will refer to the quantity which sellers will sell.

As with demand elasticity, supply elasticity depends on the amount of time available for adjustment. In the very short run there may be no adjustments they can make, which would mean a perfectly vertical supply curve. For example, if on December 1 the price of

Figure 5.1 A larger harvest can cut farm income.



apples doubles, there will be minimal effect on the number of apples available to the consumer. Producers cannot make adjustments until a new growing season begins. In the short run, producers can use their facilities more or less intensively. In the apple example, they can vary the amounts of pesticides, and the amount of labor they use to pick the apples. Finally, in the long run not only can producers change their facilities, but they can leave the industry or new producers may enter it. In our apple example, new orchards can be planted or old ones destroyed.

The topic of elasticity can help explain some situations that at first glance may seem puzzling. If American farmers all have excellent harvests, they may have a very poor year financially. They may be better off if they all have mediocre harvests. If a bus company decides it needs more revenue and tries to get it by raising fares, it may find that it in fact has less revenue than it did with the lower fares.

In the case of the farmers, the key to their problem is that the demand curve for their products is quite inelastic. This means that if the harvest is unusually good, a large drop in price is necessary to encourage consumers to use the additional grain. If the elasticity coefficient is .5, for example, and the harvest is 10% larger than the previous year, then a 20% drop in prices will occur (assuming that the many things which we keep constant in drawing the demand curve have remained constant). Because this price reduction more than offsets the effect of the larger harvest, the average farmer's income drops.

In the case of the bus company, the key is that demand is elastic. For example, suppose

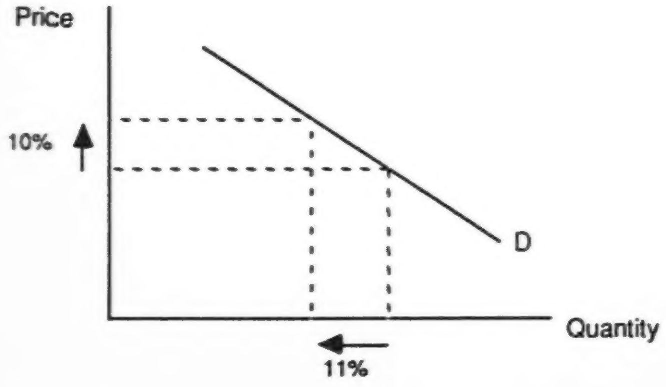


Figure 5.2 Increased bus fares can cut revenue

the elasticity is 1.1. Then if price is raised by 10%, quantity (ridership) must drop by 11%. But the drop in ridership more than offsets the increase in price, and hence revenue will drop.

COMPUTING PRICE ELASTICITY

When we try to use equation 1 to calculate elasticity coefficients, we run into a problem. If we look at an increase in price from \$3.00 to \$4.00, we have an increase of 33%. However, if we have a price reduction from \$4.00 to \$3.00, we have a reduction of 25%. Thus there are two ways of viewing the difference between what happens at \$3.00 and \$4.00.

Suppose that when price of a product is \$3.00, people will buy 60, but when price is \$4.00, they buy only 50. What is the elasticity over this segment of the demand curve, between prices \$3.00 and \$4.00? Should one start at the price of \$3.00 and compute a price rise of 33 1/3% and a quantity decline of 16 2/3%, or:

$$e = \frac{16.67\%}{33.33\%} = .5$$

Or should one start at the \$4.00 price, and compute a price decline of 25% and a quantity increase of 20%, or:

$$e = \frac{20\%}{25\%} = .8$$

The formula mentioned in the previous section does not tell us which way to proceed, and it matters. To get around the problem of deciding which starting point to use, economists compute elasticity based on the midpoint, or in the example above, at a price of \$3.50. The formula that does this is:

$$(2)\quad e = \frac{\text{Change in quantity divided by average quantity}}{\text{Change in price divided by average price}}$$

$$= \frac{\dfrac{Q1 - Q2}{(Q1 + Q2)/2}}{\dfrac{P1 - P2}{(P1 + P2)/2}}$$

Putting the numbers from the previous example into this equation yields:

$$e = \frac{(60-50)}{(60+50)/2} \Big/ \frac{(\$4-\$3)}{(\$4+\$3)/2}$$

$$= (10/55)/(1.00\ /3.50) = (10/55)\text{x}(35/10) = 7/11 = .6363\ldots$$

This formula is the formula for *arc elasticity* or the elasticity between two points on the demand curve. If one can express a demand curve as a mathematical formula, one can also, with the aid of calculus, compute *point elasticities*. We will leave this topic for more advanced courses.

With a bit of algebra, one can show that equation (2) can be rewritten as:

$$(3)\quad e = \frac{1}{\text{Slope of Demand Curve}} \times \frac{\text{Average Price}}{\text{Average Quantity}}$$

(See the question 10 at the end of the chapter.)

Using equation (3), consider what happens when the slope gets steeper, which means that the slope becomes a bigger number. (Recall that we are making price elasticity a positive number, which means we are looking at the absolute value of the slope.) Elasticity becomes smaller, which means that consumers are less responsive. As the demand curve approaches a vertical line, the slope approaches infinity, and elasticity approaches zero. As the demand curve approaches a horizontal line, the slope approaches zero, and elasticity approaches infinity.

One can also see what happens when the slope is constant, which means that the demand curve is a straight line. As one moves along the line, the elasticity will change.

Question 9 at the end of the chapter asks you to explain this.

OTHER ELASTICITIES

In addition to price elasticities of supply and demand, economists frequently refer to other elasticity measurements. **Income elasticity of demand** measures the responsiveness of people's purchases to changes in income. It is defined as:

(4) Income Elasticity = $\dfrac{\text{percentage change in amount bought}}{\text{percentage change in income}}$

Income elasticity measures whether a good is a normal or an inferior good. If the income elasticity of a product is positive, then increases in income cause an increase in the amount of that product purchased. Such a product is a normal good. For an inferior good income elasticity is negative because an increase in income causes a *decrease* in the amount bought.

Cross-price elasticity, often called just cross elasticity, measures whether goods are substitutes or complements. It looks at the response of people in buying one product when the price of another product changes. The formula for cross-price elasticity is:

(5) Cross-Price Elasticity = $\dfrac{\text{percentage change in amount of A bought}}{\text{percentage change in price of B}}$

If goods are complements, cross-price elasticity will be negative. For example, if the price of gasoline rises, the sales of large cars will decline. The positive change in the denominator (bottom) of equation 5 is matched with a negative change in the numerator (top) of the equation. The result is therefore negative. If cross-price elasticity is positive, B is a substitute for A. For example, sales of Coke will fall if the price of Pepsi falls because some Coke drinkers will switch from Coke to Pepsi.

REVENUE AND THE DEMAND CURVE

The demand curve is a tremendously useful illustration for those who can read it. We have seen that the downward slope tells us that there is an indirect relationship between price and quantity. One can also view the demand curve as separating a region in which sellers can operate from a region forbidden to them. But there is more, especially when one considers what an area on the graph represents.

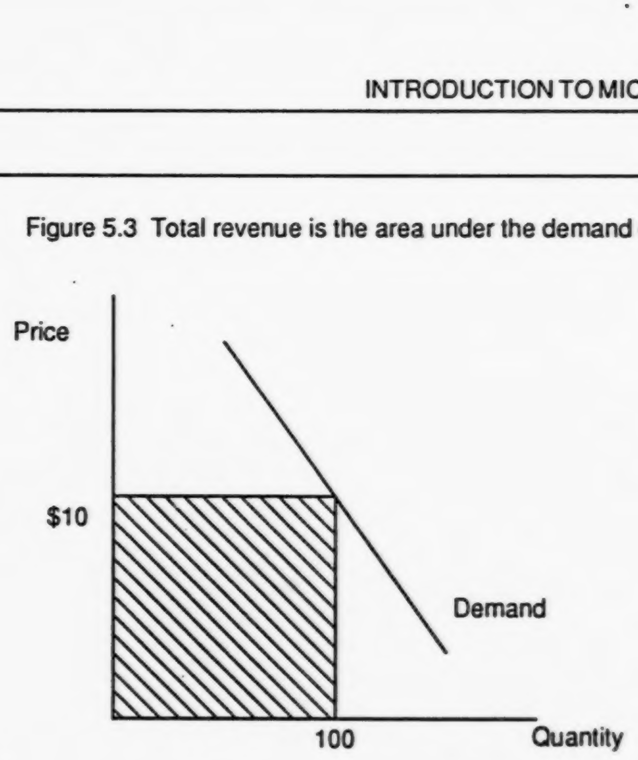
If the price of a product is $10.00 and at that price people will buy 100 units, as Figure 3 illustrates, then the total revenue to sellers, or the total expenditures of buyers, will be $1000. Total revenue equals quantity sold multiplied by price. But simple geometry tells us that if we look at the area of the rectangle formed under the demand curve in Figure 3, we can find its area by multiplying the height of the rectangle by its width. Since the height is price and the width is quantity, the area is total revenue. The fact that area on supply and demand graphs measures revenue (or expenditure, which is the same thing from another viewpoint) is a key idea which we will use repeatedly in the chapters ahead.

From the demand curve we can obtain total revenue. From total revenue we can obtain another key concept: **marginal revenue**. Marginal revenue is the additional revenue added by an additional unit of output, or in terms of a formula:

(6) Marginal Revenue = $\dfrac{\text{Change in total revenue}}{\text{Change in sales}}$

According to Figure 3, people will not buy more than 100 units at a price of $10.00. To sell more, price must drop. Suppose that to sell the 101st unit, the price must drop to $9.95. What will the marginal revenue of the 101st unit be? Or in other words, by how much will total revenue increase when the

Figure 5.3 Total revenue is the area under the demand curve.



101st unit is sold?

There is a temptation to answer this question by replying, "$9.95." A little arithmetic shows that this answer is incorrect. Total revenue when 100 are sold is $1000. When 101 are sold, total revenue is (101)x($9.95)=$1004.95. The marginal revenue of the 101st unit is only $4.95.

To see why the marginal revenue is less than price, one must understand the importance of the downward-sloping demand curve. To sell another unit, sellers must lower price on all units. They received an extra $9.95 for the 101st unit, but they lost $.05 on the 100 which they were previously selling. So the net increase in revenue was the $9.95 minus the $5.00, or $4.95.

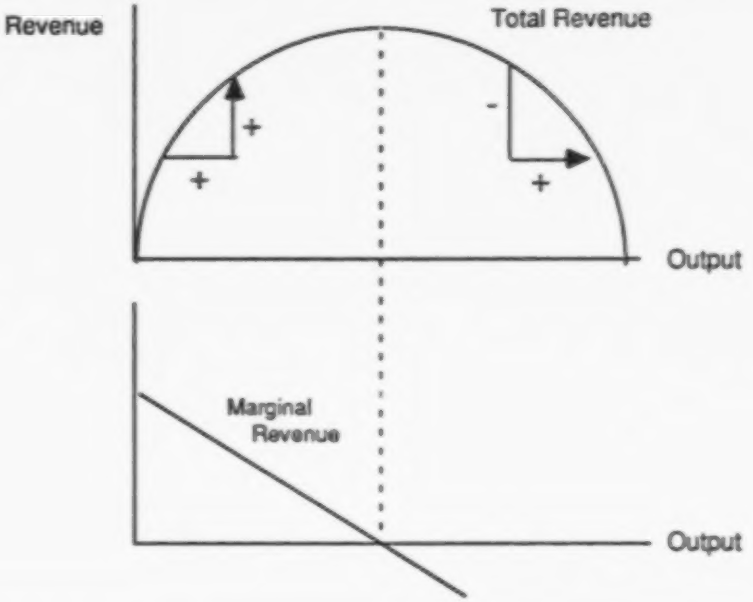
There is a another way to see why marginal revenue will be less than price when a demand curve slopes downward. Price is average revenue. If the firm sells 100 for $10.00, the average revenue for each unit is $10.00. But as sellers sells more, the average revenue (or price) drops, and this can only happen if the marginal revenue is below price, pulling the average down.

The reasoning of why marginal will be below average if average is dropping can perhaps be better seen in another example. Suppose that the average age of 20 people in a room is 25 years, and that another person enters the room. If the average age of the people rises as a result, the extra person must be older than 25. If the average age drops, the extra person must be younger than 25. If the added person is exactly 25, then the average age will not change. Whenever an average is rising, its marginal must be *above* the average, and whenever an average is falling, its marginal must be *below* the average.

If one knows marginal revenue, one can tell what happens to total revenue if sales change. If selling another unit increases total revenue, the marginal revenue must be greater than zero. If marginal revenue is less than zero, then selling another unit takes away from total revenue. If marginal revenue is zero, than selling another does not change total revenue. This relationship exists because marginal revenue measures the *slope* of the total revenue curve.

Figure 4 illustrates the relationship between total revenue and marginal revenue. The total revenue curve will be zero when nothing is sold and zero again when a great

Figure 5.4 Marginal revenue is the slope of the total revenue curve.



deal is sold at a zero price. Thus it has the shape of an inverted U. The slope of any curve is defined as the rise over the run. The rise for the total revenue curve is that change in total revenue, and the run is the change in output. Therefore,

(7) Slope of Total Revenue Curve = $\frac{\text{Change in total revenue}}{\text{Change in amount sold}}$

But this definition of slope is identical to the definition of marginal revenue in equation (6).

FROM ELASTICITY TO MARGINAL REVENUE

Marginal revenue is the extra revenue from adding another unit of output. If a firm finds that when it sells six units, its revenue is 24, and when it sells eight its revenue is 28, its extra revenue for adding two more units is four. Its marginal revenue, or the extra revenue for adding one more unit of production will be two.

Figure 5 illustrates an alternative way to compute this extra revenue. When the firm sells six, it can charge a price of 4, but when it sells eight, it can charge only $3.50. (Thus six units at $4 each gives a total revenue of $24 and eight units at $3.50 each gives a total revenue of $28.) When the firm sells the extra two units, it adds two units at $3.50 each, or $7 to its revenue. However, it also loses something because it had to lower price on the six units it was previously selling. The loss is these six units times $.5 each or $3. The net change in revenue is $7 less $3, or $4. Equation (6) says that to get marginal revenue, the change in total revenue ($4) must be divided by the change in output (2), which in this example give us $2.

We have shown that marginal revenue can be computed as

$$(8)\ \text{MR} = \frac{(\text{Change in Q})\text{P} + (\text{Change in P})\text{Q}}{\text{Change in Q}}$$

(This formula holds only approximately when changes are big, but becomes exact as the changes get very, very small. Since the change in price will be negative, the second term in the numerator will be subtracted from the first.)

When changes in price and quantity are very, very small, the formula for price elasticity can be written as

$$(9)\ e = \frac{(\text{Change in Q})}{\text{Q}} \Big/ \frac{(\text{Change in P})}{\text{P}}$$

If your algebra is fairly good, you should be able to use these formulas 8 and 9 to show that the following equation is true:

$$(10)\ \text{Marginal Revenue} = \text{Price}\,(1 - 1/|\text{elasticity}|)$$

This third formula leads to a number of interesting results, which you can derive in Question 4.

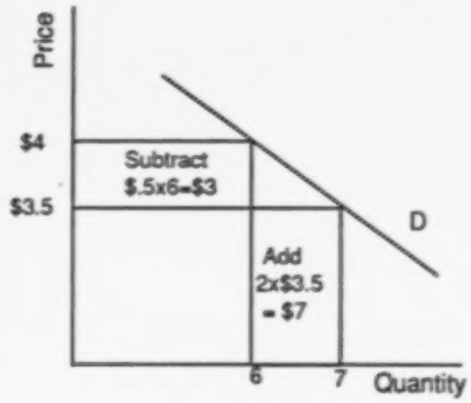


Figure 5.5 Computing marginal revenue

SUMMARY

Elasticity can be used whenever a measurable change in something causes a measurable change in behavior. This chapter discussed the most commonly used elasticity measures: price elasticities of supply and demand, income elasticity, and cross-price elasticity. It introduced you to how value can be represented on a demand-curve graph, and to the very important concept of marginal, examining how marginal, total, and average revenue are related. Finally it showed that elasticity and marginal revenue are related by means of a simple equation.

QUESTIONS

1. Use equation 2 to compute price elasticities of supply and demand in the table below. You should get three elasticity coefficients for demand and three for supply.

PRICE	Qd	Qs
$1.00	100	10
$2.00	90	40
$3.00	70	70
$4.00	40	140

2. Below is a demand curve. From it, compute total revenue and marginal revenue.

Output	Price	Total Revenue	Marginal Revenue
1	$9		
2	7		
3	5		
4	4		
5	3		

3. Below is a table showing sales and marginal revenue. Compute total revenue and average revenue.

Sales	Marginal Revenue	Total Revenue	Average Revenue
1	$12		
2	11		
3	10		
4	9		

4. (a) If the firm is a price taker, it faces a perfectly flat demand curve. This demand curve is perfectly elastic, that is, its elasticity is infinity. What is true of the marginal revenue relative to price in this case? (Hint: Use equation 10.)

(b) If elasticity is unitary elastic, that is, has a value of 1, what will be the value of marginal revenue?

(c) If elasticity is less than 1, what range of values can marginal revenue take? In this case, what will happen to total revenue if the firm sells more by cutting price?

(d) If elasticity is greater than 1, what range of values can marginal revenue take? In this case, what will happen to total revenue if the firm sells more by cutting price?

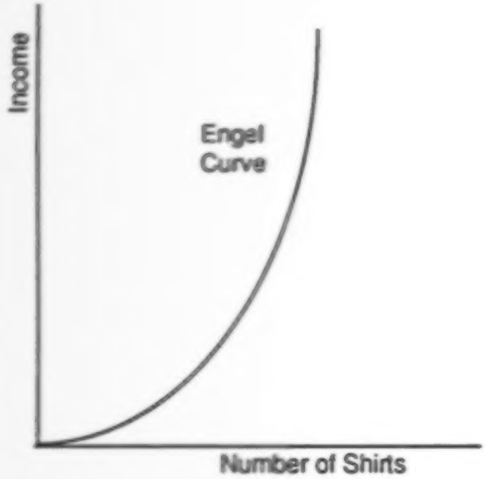
(e) Fill in the following table, putting in the words "increase," "decrease," and "stay the same" where they belong.

IF	and the firm increases sales, total revenue will:	and the firm decreases price, total revenue will:
$e < 1$		
$e = 1$		
$e > 1$		

5. When a total revenue curve reaches a maximum, what must be true of marginal revenue?

6. After six semesters of college, Greta Grouch had a grade-point average of 3.25. After her seventh semester her grade-point average dropped to 3.20. What can you tell about her grades for the seventh semester?

7. The demand curve is the relationship between the amount of a product people are willing to buy and the price of the product. An Engel curve is the relationship between the amount of a product people are willing to buy and their income. An Engel curve is shown below.



a) Does the Engel curve in the graph illustrates a inverse or direct relationship?

b) Will the income elasticity of shirts be positive or negative with this Engel curve?

c) If this curve were sloped the other way, so that the relationship were inverse, shirts would be an ______________ good.

d) If the curve were vertical, how many more shirts would a person buy if he got an extra $5000 in income?

e) A vertical Engel curve will have an elasticity of ________.

f) If the price of shirts rose, what would happen to the Engel curve?

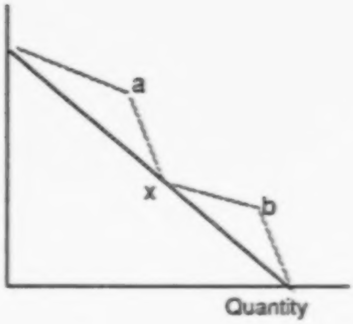
g) What is the only change which can increase the number of shirts which people will buy but which will not move the Engel curve at all?

8. If farmers face an inelastic demand curve, they as a group would be better off if each one of them cut production by 10%. Why do they not take this action? (Hint: Think of this as a prisoners' dilemma problem.)

9. (Advanced) A straight-line demand curve has a constant slope. Using equation (3), find which region (top or bottom) will be elastic and which will be inelastic.

10. (Advanced) (a) Show algebraically how one gets equation (3) from equation (2).. (Hint: begin by defining slope in terms of Q1, Q2, P1, and P2.)

(b) In a straight line demand curve, the elasticity at any point x will be length of the demand curve from X to the horizontal axis divided by the length of the demand curve from X to the vertical axis. Prove this. (Hint: Begin with equation 3 and use geometry.)


a
x
b
Quantity

CHAPTER 6

RATIONING AND ALLOCATING

After you read this chapter and work the problems at the end, you should be able to:

Explain why prices can be a way of rationing.
Explain the effects of price ceilings and price floors.
Distinguish between scarcity and a shortage.
Give examples of queuing by line and list, and coupon rationing.
Interpret a Lorenz curve.
Explain what the term "dollar voting" means.
Distinguish between positive and negative feedback.

Economics divides its study of exchange into two parts, each concerned with a different set of questions. Macroeconomics is concerned with questions of inflation and recession. Microeconomics is concerned with questions of allocation and rationing.

Allocation involves questions about ways a society can decide **what** goods to produce and **how** to produce them. Should it produce a lot of bread or a lot of clothes? Should it produce more for use now and ignore the future, or should it sacrifice consumption now and invest so that more will be available in the future? Where should it use its workers and where should it use its machinery? Rationing is concerned with the question of **for whom** are the goods and services. Who gets what is produced? Notice that these questions of allocation and rationing assume that scarcity exists—that the society cannot do everything it would like to do, but must choose some options at the expense of others.

This chapter begins with an overview of how societies can ration and allocate. Relying heavily on the model of supply and demand, it shows that rationing can be done with prices, but that there are also other ways to ration. It also shows that prices can perform as an allocation system, transmitting both the information and incentives needed to make decisions about what to produce and how to produce it. This system is sometimes called a system of "dollar voting." The chapter also describes an alternative way to make allocation decisions, the way of central planning.

PRICE AS A RATIONER

There have been many cases throughout history in which governments have for some reason been unwilling to let markets adjust to market-clearing prices. Instead they have established either **price ceilings**, which are prices above which it is illegal to buy or sell, or **price floors**, which are prices below which it is illegal to buy or sell.

If a price ceiling is placed below the market-clearing price, as Pc is in Figure 1, the market-clearing price becomes illegal. At the ceiling price, buyers want to buy more than sellers will make available. In Figure 1 buyers would like to buy amount Q4 at price Pc, but sellers will only sell Q1. Because they cannot buy as much as they would like at the legal price, buyers will be out of equilibrium. The normal adjustment that this would set into motion in a free market, an increase in price, is illegal, and either buyers or sellers or both will be penalized if transactions take place above Pc. Buyers are faced with the problem that they want to buy more than is available. This is a rationing problem.

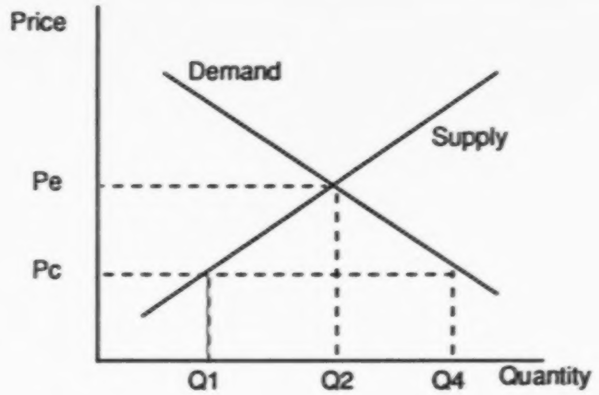
There are two commonly used ways to solve the rationing problem caused by price ceilings. One is by **queues**. A queue is a line, and it solves the rationing problem by a "first come, first served" solution. Though price ceilings fix the monetary cost that the buyers must pay so that buyer equilibrium cannot be restored by higher prices, they do not fix the nonmonetary cost of waiting.

Under a system of queuing by waiting line, waiting time is the variable that changes to restore buyer equilibrium. A person who is willing to buy 5 of an item for $1.00 when there is no waiting time may not be willing to buy any at all if the price is $1.00 and waiting time is two hours. With a system of queuing by line, waiting time will rise until enough buyers have dropped out so that there is a match between the amount available and the amount people are willing to buy.

Queuing is a rarity in the United States, but in many parts of the world, especially in Eastern Europe and the Soviet Union, it is a part of daily life. Hedrick Smith explains queuing in these words:

> "The only real taste of stoical shopping vigils in recent American history were the pre-dawn lines at service stations during the gasoline crisis in the winter of 1973-4.... But it was temporary and only for one item. Imagine it across the board, all the time, and you realize that Soviet shopping is like a year-round Christmas rush. The accepted norm is that the Soviet woman daily spends two hours in line, seven days

Figure 6.1 A price ceiling makes high prices illegal.



> a week. . . . I noted in the Soviet press that Russians spend 30 billion man-hours in line annually to make purchases.... 30 billion man-hours alone is enough to keep 15 million workers busy year-round on a 40-hour week."[1]

Smith notes that coping with shortages and queuing requires a different approach to shopping than that with which Americans are familiar. People tend to carry large amounts of cash and a bag or briefcase whenever they go out just in case they find some desirable good. When they see a line, they join it because there must be something worthwhile at the front. People shop not just for themselves but for friends and kin, and as a result "know by heart the shoe, bra, pant and dress size, waist and length measurements, color preferences and other vital particulars for a whole stable of their nearest and dearest...."[2]

Chapter 1 noted that it was possible to explain the gasoline shortage and the resulting queuing in the early 1970s in terms of a good-vs-bad model, but that economists did not use this model. Rather economists consider the shortages the direct result of price controls in existence at the time. A move by certain oil-producing nations caused the supply curve for gasoline to shift to the left. When U.S. government price restrictions kept pump prices from rising to equilibrium, shortages resulted. These shortages were predictable because in this model price is a rationing device. In contrast, the only function of price in a good-vs-bad model is to take money from one group and give it to another.

A second system of nonprice rationing is with coupons. In this system the government distributes coupons that must be presented along with money in order to buy a product. Coupons restore buyer equilibrium because they change the cost of a product, though in a different way than queuing does. Under the system of queuing, the cost of a product is its price plus waiting time, and under a system of coupon rationing the cost of a product is its price in money plus its price in coupons. The United States had an extensive system of coupon rationing during World War II. Figures 3 and 4 show coupons and booklets issued during the Second World War.

The existence of price ceilings is often accompanied by illegal transactions, or a **black market**. Black markets exist when

---

[1] *The Russians* (Quadrangle/The New York Times Book Co., 1976), p. 64.

[2] *Ibid.*, p 62.

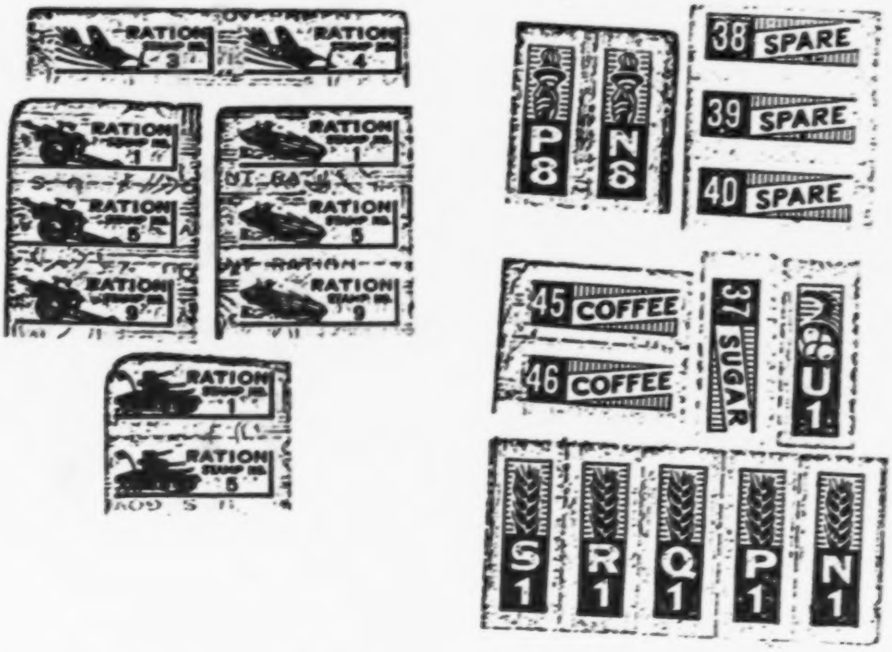


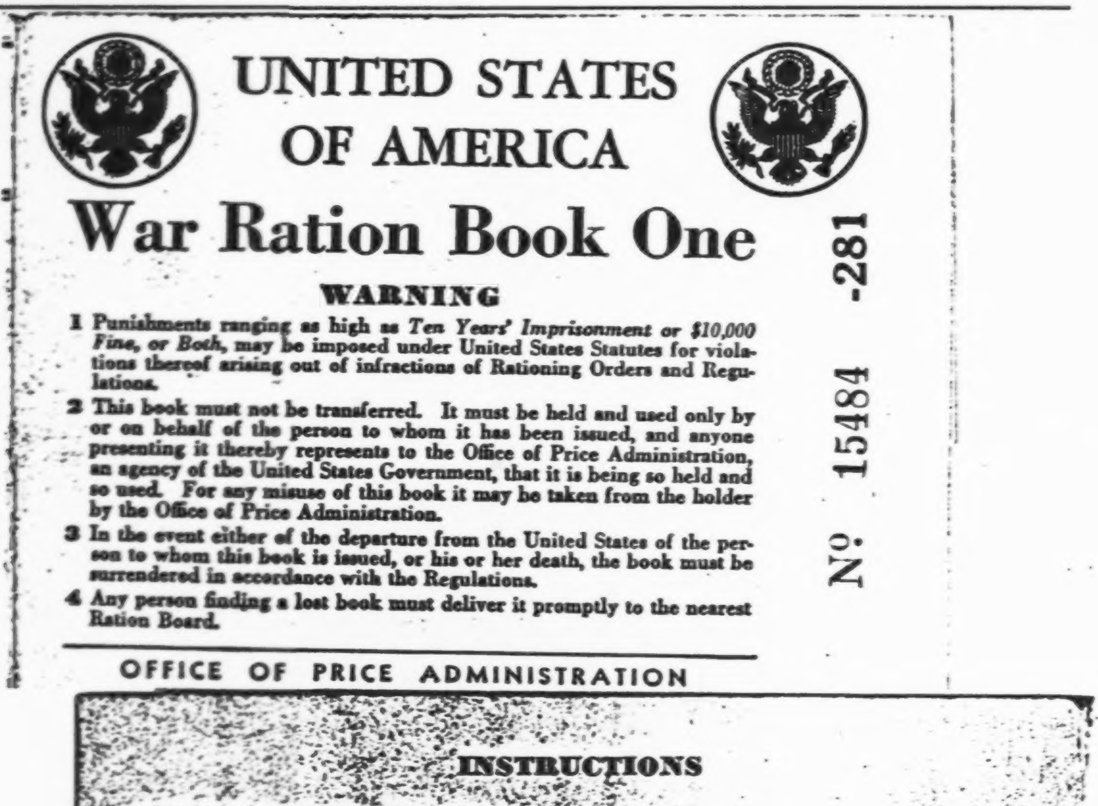
Figure 6.2 Ration stamps from World War II--from book 3 (left) and book 4 (right).

people try to "beat the system," either the system of coupons or queuing. In an economy of free markets there can be no black markets.

Price ceilings are not the only sort of price controls governments have imposed. There have also been many laws which establish *minimum* prices, or price floors. Figure 4 illustrates a price floor with price Pf. At this price buyers are in equilibrium, but sellers are not. They would like to sell quantity Q2, but buyers are only willing to take Q3. To prevent the adjustment process from causing price to fall, government may buy the surplus as the U.S. government has done in agriculture and in precious metals. If it does not buy the surplus, the government will have to penalize either buyers or sellers or both who transact below the price floor, or else price will fall. Since there is no one else to absorb the surplus, sellers will.

Rationing is necessary to deal with scarcity. When an item is scarce, people must sacrifice something in order to get as much of the item as they would like to have. There are some goods which are not scarce. Air is an example—it is free to all who want to breathe it. Ice is not scarce in Greenland. But almost all other goods are scarce. Price is a way to ration goods. It deprives those who do not have enough income or desire for a product. The function of price as a rationer is most clearly seen when price is prohibited from acting as a rationer, so that some other method of rationing (such as queuing or coupon rationing) must emerge or be developed.

Note that scarcity is not the same as shortage. Scarcity exists whenever some rationing process must be used. A shortage indicates that price is below market-clearing price.



UNITED STATES OF AMERICA

War Ration Book One

Nº 15484 -281

WARNING

1 Punishments ranging as high as *Ten Years' Imprisonment or $10,000 Fine, or Both,* may be imposed under United States Statutes for violations thereof arising out of infractions of Rationing Orders and Regulations.

2 This book must not be transferred. It must be held and used only by or on behalf of the person to whom it has been issued, and anyone presenting it thereby represents to the Office of Price Administration, an agency of the United States Government, that it is being so held and so used. For any misuse of this book it may be taken from the holder by the Office of Price Administration.

3 In the event either of the departure from the United States of the person to whom this book is issued, or his or her death, the book must be surrendered in accordance with the Regulations.

4 Any person finding a lost book must deliver it promptly to the nearest Ration Board.

OFFICE OF PRICE ADMINISTRATION

INSTRUCTIONS

1 This book is valuable. Do not lose it.

2 Each stamp authorizes you to purchase rationed goods in the quantities and at the times designated by the Office of Price Administration. Without the stamps you will be unable to purchase those goods.

3 Detailed instructions concerning the use of the book and the stamps will be issued. Watch for those instructions so that you will know how to use your book and stamps. Your Local War Price and Rationing Board can give you full information.

4 Do not throw this book away when all of the stamps have been used, or when the time for their use has expired. You may be required to present this book when you apply for subsequent books.

Rationing is a vital part of your country's war effort. Any attempt to violate the rules is an effort to deny someone his share and will create hardship and help the enemy.

This book is your Government's assurance of your right to buy your fair share of certain goods made scarce by war. Price ceilings have also been established for your protection. Dealers must post these prices conspicuously. Don't pay more.

Give your whole support to rationing and thereby conserve our vital goods. Be guided by the rule:

"If you don't need it, DON'T BUY IT."

10—62300-1 ☆ U. S. GOVERNMENT PRINTING OFFICE : 1942

Figure 6.3 Coupon books from World War II--the front cover of book 1 (top) and the back of book 2 (bottom).

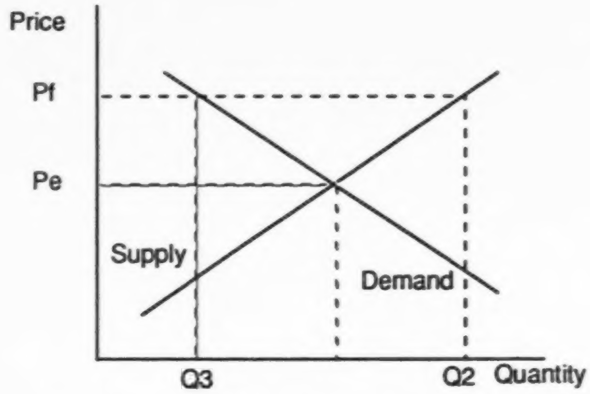


Figure 6.4 A price floor makes low prices illegal

Price ceilings are usually justified as a way to help consumers, but whether they actually do is open to doubt. Consider, for example, rent controls, a popular form of a price ceiling. If the demand curve and the short-run supply curves are inelastic, then a sizable drop in rents may result in a very small shortage. The benefits to consumers will, in the judgement of most, clearly outweigh the costs to the consumers. Further, the short-run supply of housing should be quite inelastic because apartment buildings take time to build and even longer to wear out.

But apartment buildings do wear out, and they wear out much faster when they are not properly maintained. Effective rent controls will discourage the construction of new buildings and encourage the retirement of old buildings. With time, sellers will approach the long-run supply curve, and the small initial shortage may become increasingly large. Rationing will be on a first-come, first-served basis, and under-the-table payments will be encouraged. In the long run the costs to consumers may considerably outweigh the benefits. Even if this is the case, however, the program may remain politically popular if those who benefit from the program, those living in rent-controlled apartments, vote while those harmed are forced to live in another political jurisdiction and thus will not be able to vote against the program.

New York City has had a system of rent controls since World War II. It is a complex program, with not all rental property controlled. Yet it has had consequences which a supply and demand model leads one to expect. It is very difficult to obtain an apartment in a rent-controlled building. Also, the city has had a major problem with property abandonment. Those planning to abandon (which is illegal) try to maximize their cash flow by cutting or eliminating maintenance and by not paying taxes.

The most visible price floor in the United States is the minimum wage. The U.S. Congress passed a minimum wage law in 1938 and has raised its level and extended its coverage several times since then. The stated goal of the minimum wage is to help the poor. It will not directly affect most workers because they have wages which are above the minimum. Only those workers who are earning less than the minimum will be directly affected.

Figure 4 indicates that the minimum wage will help some people. Some people whose wages are below the minimum will see their wages rise. But others will be harmed. Some tion and distribution of wealth were distinct subjects, and "[t]he distribution of wealth, therefore, depends on the laws and customs of society."[2]

Two things are worth noting at this point. First, Mill is using the word "laws" in two distinct ways in these quotations. In the first use he is using it as meaning a regularity, and in the second he is talking about legal requirements. Second, most economists have decided Mill was wrong that production and distribution can be separated. They rather are two sides of a single coin.

Mill spent two chapters discussing how to remedy low wages. His final solution was in the tradition of Malthus and Ricardo. He believed that the problem would go away if the poor had fewer children.

The question of poverty was at least as interesting to Alfred Marshall. At the very beginning of his *Principles*, he states that the question of whether poverty may be extinguished "gives to economic studies their chief and their highest interest." Yet both Marshall and Mill had a different emphasis than that which is dominant in the second half of the twentieth century. They saw the elimination of poverty as a process of economic growth. Many today see it as a process of redistribution, that is, taking from the rich and giving to the poor through taxation and transfers.

It is possible to measure how equally or unequally the price system rations by looking at the distribution of income. Table 1 shows what these results looked like in the United States in 1978. It shows that 20% of households (groups of people living together, usually families, or single people if they live alone) had total money incomes of less than $6,391. These people received only 4.3% of the total income which households earned. Twenty percent of households earned between $6,391 and $11,955, and these households earned 10.3% of the total income earned. The rest of the table can be interpreted in the same way.

If the information on the table is put onto a graph as in Figure 5, a **Lorenz curve** is produced. The further the Lorenz curve lies below the line of equality, the more unequal is the distribution of income.

All economic statistics have problems, and the Lorenz curve and the numbers from which it is constructed are no exceptions. Problems come from two sources. First there is the question of whether the number actually measures what it is supposed to measure, and second are the difficulties in making the measurement.

Income distribution is intended to tell us about the rich and the poor, or about how much discrimination exists in a system of price rationing. In a system of price rationing, however, differences in the ability to use income wisely also determines how much discrimination there is. If those who receive the most income, for example, also tend to be the most capable at using that income, then the picture that the Lorenz curve shows will understate the actual amount of inequality.

If rationing is not done solely by price, but by other methods as well, then looking at income data may be meaningless. In the United States most rationing is done with price, but not all. For example, the purpose of public housing and food stamps is to prevent rationing by price. Both of these items are ignored in the data in Table 1. Also, great caution must be used in comparing income distributions among countries because their rationing systems can be very different. For example, comparisons of income distribution between the United States and the Soviet Union are probably not meaningful—even though economists sometimes make them—because the Soviet Union not only relies heavily on queuing, but those with special status, such as party members, have access to stores denied the ordinary citizen.

Households differ in size and average age, but these differences are not reflected in Table 1. Neither is the fact that the amount of time over which income is earned affects the shape of the Lorenz curve. Larger house-people will not be able to find work at the new, higher wage. They will not be hired if the value of the work they can do is less than the minimum wage. These people will become the surplus which a price floor generates. They, however, may not realize the cause of their difficulties, and so may not realize that the minimum wage harms them.

---

[2] *Ibid.* p. 197.

[3] *Principles of Economics*, 8th ed. (New York: The Macmillan Company, 1920), p. 4.

Many advocates of minimum wages do not argue in terms of the model of supply and demand (though some do). Rather they use a power model which sees the issue as a conflict between employer and employee. If employers have power, then they may pay less than the value that is or can be contributed by the worker. In a sense the employer "exploits" the worker. In this model the minimum wage helps workers at the expense of employers' profits. In contrast, the model of supply and demand suggests that the minimum wage helps some workers at the expense of others. Obviously it is more difficult to argue in favor of a minimum wage using the model of supply and demand.

The question of which model is the better model is an empirical question, that is, one settled by an examination and analysis (in this case a statistical analysis) of the facts that are available. Economists have done numerous studies trying to find out what the effects of the minimum wage are. The vast majority suggest that the minimum wage does have some adverse employment results. They find that the minimum wage results in unemployment for some, especially those whose skills and abilities are very low, and higher wages for others. This is what the model of supply and demand predicts will happen.

RATIONING AND THE DISTRIBUTION OF INCOME

The previous section showed that prices were a way of rationing goods and services. It should be obvious that not everyone does equally well under a system of price rationing. The system discriminates in favor of some and gives them large amounts of goods and services, and against others to whom it gives very little. We call the first group the rich and the second group the poor.

People with skill and ability, luck, and willingness to work hard will do well under a system of price rationing. They will be able to either sell their labor to others for high amounts, or will themselves be able to produce things (goods, services, or ideas) that others value highly. Some who do well have had the luck to inherit wealth, usually from an ancestor who had skill, ability, and willingness to work hard. There are others who do well because they evade the rules of the system. They use deceit and force to obtain wealth and income. The poor, on the other hand, are those who for some reason do not have anything to sell that others value highly.

Income is not the only item that determines how well people do under a system of price rationing. Some people know how to use their money better than others. Some people can "stretch" the dollar, or "make it go a long way." People must pursue their goals in a world of uncertainty and partial knowledge. Those who can distinguish a product of high quality from one of low quality, who can buy when and where prices are lower, and who buy only those items that best lead to their goals, will do better under a system of price rationing than those who do not have these abilities.

Interest in the extent to which a price system discriminates for and against people extends back to the early economists and continues today. John Stuart Mill reflects the views of the classical economists (who were the dominant school of economic thought for a century after 1776). In his *Principles of Political Economy*, first published in 1848, he explains: "Writers on Political Economy profess to teach, or to investigate, the nature of Wealth, and the laws of its production and distribution."[1] Mill decided that the produc-

---

[1] Volume 1 (New York: The Colonial Press, 1900), p. 1.

TABLE 1
PERCENT DISTRIBUTION OF AGGREGATE HOUSEHOLD INCOME IN 1978, BY FIFTHS OF HOUSEHOLDS

Households	Percent of Income
Lowest fifth (under $6,391)	4.3
Second fifth ($6,392 - $11,955)	10.3
Third fifth ($11,956 - $18,122)	16.9
Fourth fifth ($18,122 - $26,334)	24.7
Top fifth ($26,335 and over)	43.9

Source: U.S. Bureau of the Census, *Current Population Reports*, P-60, No. 121, "Money Income in 1978 of Households in the United States," U.S. Government Printing Office, Washington, D.C. 1980. Data taken from cover. (Data are before taxes.)

holds tend to earn more than smaller households. People in their thirties tend to earn more than people in their twenties. Households with four or five members, with more than one person working, and whose working members are between 35 and 55 tend to earn more than other households. In a paper published in the *American Economic Review* in September of 1975, Morton Paglin concluded that ignoring the influence of age on earnings overstates inequality by 50%. There is also variability from year to year in how much households earn. Some people appear poor because they had an unusually bad year, and others will seem rich because they had an unusually good year. The shorter the period over which income is measured, the more unequal the distribution appears. Thus, if income were measured over a decade, the distribution would be more equal than any of the yearly distributions.

The other source of problems is in making the initial measurements. The data shown in Table 1 were obtained from questionnaires given a sample of 56,000 households. Not all of these households gave correct answers. The publication containing these data had a lengthy discussion of measurement problems, but when other people use these data in a book or an article or an argument, that lengthy discussion often gets left out (as it does here).

The problem with measuring the amount of discrimination in a system of price rationing (and you should be aware that any rationing system must discriminate—it must decide who gets each product and who does not) are sizable, but not unusual when it comes to economic measurement. There are also substantial problems in the most common of economic statistics, such as the unemployment rate, the rate of inflation, and the gross national product.

Even though there are measuring problems, it is clear that a system of price rationing will distribute goods less equally than will alternative systems such as those using queuing or coupons. Many people consider this inequality a major shortcoming of a market economy, and most critics of market systems emphasize this characteristic. De-

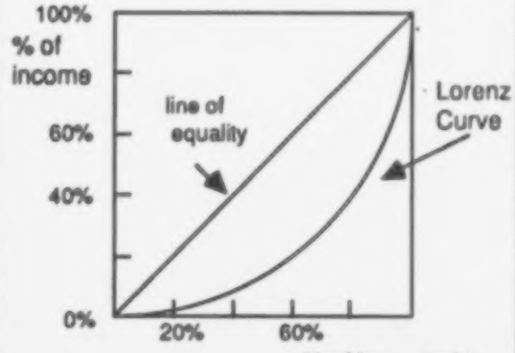


Figure 6.5 A Lorenz curve illustrates inequality.

fenders of market systems, on the other hand, tend to ignore rationing issues, and instead focus on the ability of a market economy to coordinate information and incentives. These are tasks which markets seem to perform very well in comparison to the ways other systems do them.

PRICES AS INCENTIVES AND A COMMUNICATION NETWORK

Prices are a method of rationing, a rationing system that needs no other. A previous section indicated that queuing and coupon rationing could also be used to ration goods. But it is doubtful that they can stand alone without any prices or money. The problem is not their ability to ration; rather it is that prices and money provide an important incentive for people. Most supply curves will show zero quantity at a zero price.

It may be possible to solve the incentive problem with coercion; studies of slavery indicate it was a profitable system in the old South, and studies of the concentration camps of Nazi Germany and Stalinist Russia will probably show that they too were profitable—that is, they produced more than it took to maintain the camps. The point here is that if a system tries to eliminate money and prices entirely, it will have to find not only an alternative method of rationing, but an alternative method of giving incentives. There is good reason to believe that alternative methods of incentive will be morally repugnant to many people.

There is still another problem of doing away with price and money. They not only give incentives, but they also provide a convenient way to summarize information to decision makers about wants and availability of goods. To see this, consider a commodity, say copper. The supply and demand model may only approximate this market because there are only a few producers, but the approximation is probably fairly good.

Suppose a new use is developed for copper which will use 10% of the available supply. Prices will communicate this information to all copper users and producers. The demand curve will move to the right and price will rise, telling users that copper now has increased value elsewhere and that they should try to cut back. The rise in price will tell producers that they should try to produce more, and will reward them if they do. Hence if an important new use for copper is developed, it is not necessary for either producers or users to know what it is. Prices give them all the information they need to know to conduct their affairs.

Or suppose that a major new source of copper is found. This will cause the supply curve to shift right and price will fall. The message that price gives users is that they should consider using more. The lower price tells them that the resource is now more abundant relative to its uses. The message the new price gives to producers is that they should consider reducing their efforts to produce copper, since their efforts are now of less value, and they should consider shifting resources into other lines of production. There is again no need for any producer or consumer to know the details of the new supply. Prices tell them all they need to know to make correct decisions.

The role price plays as a communication network is best seen when it is suppressed, just as the role of prices as a rationing system is best seen when it is suppressed. The 20th century has witnessed economies that have largely suppressed it. When the Bolsheviks triumphed in the Russian Revolution of 1917, they proceeded to destroy the old system which was in part a market system. Since Karl Marx had studied only capitalism, and his reason for studying it was to explain why it should not be able to succeed economically, the new rulers of Russia could not look to him for an analysis of how socialism would work. They had to learn for themselves. They quickly learned that if decisions were not made using prices and profit, another method had to be used. They developed a system called central planning. (Recall that Chapter 3 introduced central planning as one of two ways a society could coordinate actions.)

With a system of central planning, the government decides what the priorities of the society will be and implements these priorities through a central-planning agency. However, central planners cannot simply issue orders and expect them to be obeyed. They must first obtain information about what the factories, mines, and farms in the nation are capable of producing. The central planning agency could send out questionnaires, but there is no assurance that the answers that would come back would be truthful, because despite the socialist dream that men should act in total selflessness—for the good of the society as a whole—this does not seem to be the way that people actually act in the real world. It certainly has not been the way that men have acted in the nations that have adopted socialism.

The central planning agency wants the factories, mines, and farms to produce as much as possible. It must reward good performance and penalize poor performance. Good performance involves producing up to potential. This introduces an incentive to lie about potential when the central-planning agency wants to know potential. If the plant, mine, or farm manager underestimates potential, and the central planners believe this estimate, the chances that the production unit will appear to be a good performer are enhanced. Since profit is not used as basis of judging good performance (it seems too capitalistic), the central planning system in the Soviet Union judges performance on how well the productive unit meets a set of goals specified in terms of quantities. Managers have an incentive to withhold information about potential and to overestimate the supplies needed to produce their output.

Once the central planners have information about what factories are capable of producing, they must coordinate production decisions. One factory's output is often another factory's input. In planning for truck output, did the planners plan for enough steel? To get the steel, did they plan for enough machines to mine the coal and enough trucks to transport it? In a modern economy there are a tremendous number of linkages and interdependencies. Thus any change in production of one product sets off a cascade of other changes needed to support the first change. The Soviet Union has shown that a country can centrally plan its economy, but their experience has also shown that such a system tends to be rigid and resistant to change.

Between 1920 and 1940 a number of economists debated about whether a centrally-planned economy could match the performance of a decentralized economy. This debate outlined the problems which a centralized economy would have using information were outlined Friedrich Hayek argued that information is widely scattered in society and cannot be effectively collected for use by a central authority making production decisions for the entire economy. Rather the existence of market-determined prices communicates vital information which encourages individuals not only to use whatever information they have about production of goods, availability of resources, and how to satisfy consumers' desires, but also to actively search for more information. The market, said Hayek, was a way a society minimized effort needed to discover and communicate information. The, unimpressive performances of centrally-planned economies in providing for the average consumer's material well-being seems to be mostly explainable in terms of problems of information and incentive.

The fact that prices carry not only information but incentives for economic agents to act means that prices are important in deciding what goods are produced, or in the jargon of economists, they help allocate resources. Consumer preferences, the relative availability of resources, and technology determine how much of each good and service will be produced in a market system.

Economists sometimes make an analogy to politics to explain how a price system decides what should be produced. A political election offers citizens a variety of candidates, each of whom tries to convince the citizens that his policies are best. In the process of voting, the electorate keeps some candidates and retires others to private life. Because political leaders must keep coming before the citizens for election, the actions they can take are limited. The electorate has ultimate control over the policies which a democracy adopts.

Citizens, as consumers, also vote in the marketplace. The marketplace offers consumers a variety of products, and each producer tries to convince consumers that his product is best. The consumers vote for products by spending their dollars. Those products receiving a lot of **dollar votes** will be profitable and will continue being produced. Those products which do not receive enough dollar votes will die. Dollar voting provides feedback to producers. It tells them whether their performance is acceptable or not. Dollar voting implies that consumers, not producers, ultimately decide what will be produced in a market economy. This power of consumers is often referred to as "consumer sovereignty."

People who like the market system tend to stress its responsiveness to consumer desires. Compared to systems of central planning, markets usually perform well on this basis. Later chapters, however, will show that there are conditions under which markets falter and consumer sovereignty is weakened. Also, there are times in which a society may not want the consumer to be sovereign. It may decide that there is some social goal more important than individual goals. When nations are at war, for example, they usually want to lessen consumer sovereignty. As a result, in times of war most nations reduce the importance of market decisions and increase the importance of allocation by government direction and regulation.

PRICE AND FEEDBACK

A way to view the functions of price which the previous section has explained is in terms of the concept of feedback. A **system of feedback** exists when two variables are interrelated, so that a change in variable A affects variable B, but in turn a change in variable B affects variable A.

Biological organisms and systems are full of feedback mechanisms. One simple one which you understand involves body temperature (variable A) and the amount a person sweats (variable B). The higher body temperature, the more a person sweats. However, the more a person sweats, the lower will be body temperature. In this system body temperature causes sweating, but sweating in turn causes body temperature. Feedback in this case is stabilizing, and is an example of **negative** feedback. Negative feedback is illustrated in Figure 6.

Feedback can also be destabilizing. For example, as land erodes, it supports less vegetation. But land with less vegetation erodes more readily. This feedback loop is **positive** feedback, and is illustrated in Figure 7.

Feedback systems occur all around us. (See question 7 for examples.) The ones that are the most important in economics use the negative feedback systems, or stabilizing systems, of markets. If a product becomes less available, its price rises. This rise in price causes a number of actions to occur which limit the rise, and may in some cases set the stage for a future reduction in price.

First, higher prices cause people to use less of a product. They may substitute or do without. Higher prices of petroleum in the 1970s, for example, encouraged people to drive less, to work closer to the place they lived, to insulate their homes, to turn down thermostats in winter and to turn up thermostats in summer, to buy smaller and more fuel-efficient cars, and to find a host of other ways to cut usage. The higher the price, the greater the incentive for consumers to buy less.

Second, higher prices encourage producers and potential producers to find ways to produce more of the product. As a response to the oil price increases, producers found ways to extract more oil from old wells. They also found it worthwhile to greatly step up the search for new oil fields around the world, and a considerable amount of new oil was found as a result of this search.

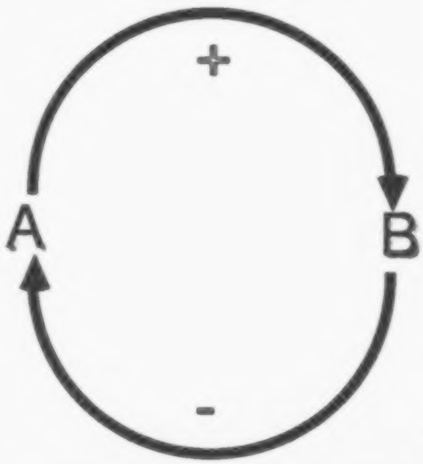
Third, higher prices signal people to search for new ways to meet old demands. The higher prices of oil spurred attempts to increase the energy efficiency of machines. It encouraged people to switch to substitute fuels. It spurred research on alternative energy sources, including new ways to tap the energy in solar radiation, wind, and water.

The negative feedback system in a market also works when the price falls. When a product becomes cheaper, reactions opposite to those described above are set into motion. When oil became cheaper in the mid 1980s, people were encouraged to use more, producers were encouraged to search for less, and everyone was encouraged to stop substituting away from oil.

When viewed in terms of the concept of feedback, the market system can be seen as a system that has important self-stabilizing features. Sometimes these features are ignored. For example, in the early 1970s there were some highly publicized "doomsday" forecasts which foresaw a rapid depletion of the earth's resources and a collapse of industrial societies. These studies often ignored price effects entirely, simply assuming that one could extrapolate present usage trends into the future. In this view, the use of the resource would continue to rise until suddenly it would be depleted, and usage would collapse to zero.

The future is very hard to predict with any accuracy. Perhaps someday industrial civilization will collapse because of resource depletion. But anyone who believes such a forecast of the future should be able to explain why this outcome is reasonable given the negative feedback system of price incentives. If they ignore this feedback, they must find market systems unstable, but they do not deserve serious intellectual consideration.

Figure 6.6 Negative Feedback



Feedback systems can have delays in adjusting. A system with negative feedback does not have to be in equilibrium, but it does have forces which keep it close to equilibrium. In the next ten chapters we will not pay

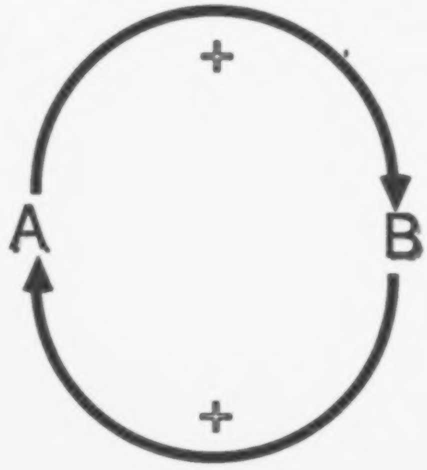


Figure 6.7 Positive Feedback

much attention to the question of adjustment. Only when we discuss inflation and unemployment in volume 2 will we return to this issue of feedback and adjustment.

SUMMARY

This chapter considered the ways societies can decide what to produce and for whom to produce. In looking at the *for-whom* question, it examined three methods of rationing: by price, by queue, and by coupon. It then discussed how equally a price system rations by looking at the distribution of income. Finally, it briefly compared two ways of answering the *what* question, the market and central planning.

QUESTIONS

1. What will the effects of a price ceiling be if it is set above the market-clearing price?

2. A government bureaucrat wants to set a price in a market so that the amount bought and sold is maximized. He asks you whether he should impose a price ceiling or a price floor. Assuming that the market can be described by supply and demand curves, what should you tell him? Explain.

3. Using the data below, construct a table similar to Table 1. Determine whether income is distributed more or less equally in your table than in Table 1. (Hint: Divide the households into five groups, (the top earners, the lowest earners, etc.) and find the percentage of income going to each group.)

Household	Income
A	$80,000
B	$40,000
C	$37,000
D	$30,000
E	$28,000
F	$24,000
G	$18,000
H	$16,000
I	$14,000
J	$13,000

4. In the late 1970s gasoline prices were rising rapidly and there was considerable public support for proposals to set a price ceiling and issue ration coupons. Supporters of rationing wanted to ration gasoline on the basis of "need," but were not in agreement on how need could be measured for actual distribution of coupons. Some bases for distribution include the number of cars a household has, the number of licensed drivers a household has, equally to all households or persons, and the number of miles each household must drive to work. Explain why some people would consider each of these bases unfair. Can you think of a method that you believe is clearly superior to any of the above methods?

5. Suppose a society switches from a system of price rationing to a system relying heavily on queuing.

a) Which of these households would benefit from and which would be harmed by the switch: families which earn large incomes by working long hours, families with small children and no friends, the unemployed, the healthy retired, families in which both husband and wife work.

b) Will a system of queuing ration more or less equally than a system of prices?

c) Will this switch encourage or discourage production?

d) The elite in the Soviet Union do not wait in queues, but shop in special stores. Why is the establishment of special stores almost inevitable in a system of queuing?

6. There are other ways to ration than those mentioned in the chapter. A lottery is one way, and queuing by list is another. Many libraries use queuing by list for new and popular books. For example, if your college library gets *How to Master Economics with Absolutely No Effort*, and fifty students want to read the book, the library may establish a waiting list. Under this system of queuing, people do not have to waste time standing in line, but it is unlikely that the waiting list will be fifty students long, even though it is almost costless to add one's name to the list. The date on which a student gets the book will matter to most students. Reading the book after one graduates, or after one has finished economics courses, may have no value. Buyer equilibrium is also restored in queuing by list with time, but here waiting time decreases the value of the good being rationed.

a) Suppose the list is 16 students long. The library decides to cut the list in half by obtaining a second copy of the book. Will it succeed in cutting the list in half? Explain.

b) Suppose that instead of maintaining a list, the library re-shelves the book each time it is returned. What forms of behavior on the part of students will develop as they try to get the book?

7. Decide whether each of these cases of feedback below represents positive or negative feedback.

a. The temperature rises as the furnace runs; as the temperature rises, the furnace turns off.

b. A teacher praises a child when she performs well; the child performs well when she is praised.

c. A valve releases more steam as the pressure rises; the pressure drops as steam is released.

d. When less gasoline is available, the price rises; when price rises, people use less gasoline.

e. When Joe studies more, he gets higher test scores; as he gets higher test scores, he studies less.

f. When wood burns, it creates heat; when heated, wood burns more readily.

g. When income goes up, people spend more; when people spend more, income goes up.

8. (a) Cities in which a large amount of housing is owned by those who use it tend not to have rent controls. Why is it not in the interests of homeowners to support rent control? (Hint: Consider taxes—those raised from rental property and those raised from homeowners.)

(b) The District of Columbia initiated rent controls in 1974 as a temporary measure intended to last until vacancy rates rose. Those establishing the controls argued that high vacancy rates would keep rents low. But until these vacancy rates rose, the government would be needed to keep rents low. What prediction would an economist make about how long Washington will have rent controls, and why?

CHAPTER 7

THE LOGIC OF CHOICE

After you read this chapter and work the problems at the end, you should be able to:

Define utility, tradeoff, cost, and law of diminishing marginal utility.
Given income and prices, be able to form a budget constraint.
Explain what a production-possibilities frontier shows, and what factors can move it.
If given a table of total costs or benefits, be able to compute marginal costs and benefits, and vice versa.
When given a table or graph showing costs and benefits for various levels of activity, use the maximization principle to find optimum positions.
When given an income and a table showing marginal utilities and prices of two items, use the equimarginal principle to find optimum positions.
Explain why the marginal cost curve measures the slope of the total cost curve.

By this time you should be familiar with the idea that people make decisions by weighing costs and benefits. This idea can be stretched to explain the amounts of goods which people buy or sell. The key is to see a decision about the amount to buy or sell as a series of small decisions. To decide how much to buy, for example, a person should consider the costs and benefits of buying the first item, then the costs and benefits of buying the second, etc. He then buys in all cases in which the benefits outweigh the costs. This idea is simple and vital in economics. The complications arise when we set out to determine costs and benefits.

This chapter begins by looking at goals as the source of benefits and constraints of scarcity as the source of costs. It then shows how one can arrange information obtained from goals and constraints so that one can make decisions based on costs and benefits. Finally, it shows some applications which economists have made to the techniques which the chapter explains.

GOALS AND BENEFITS

Economists assume that people have goals and work to obtain those goals. Economists are vague about what the goals of people are, though they expect that material advancement (for self or family) is important for many people, and they say nothing at all about the source of goals or their desirability. Nonetheless, the assumption that people strive to obtain their goals as best they can, given limitations that the world imposes on them, forms the heart of virtually all economic theory.

Goals are complex and often a bit fuzzy. Often it is hard to see exactly what goal a person has. For example, there are many people who smoke and yet say they want to quit. Such people are at war with themselves. One part pulls them one way, and another part pulls them in an opposite direction. In addition, people often do not know what it is that they want, or if they do know what they want, they do not always know how to obtain it. Most of us want to be happy, but many of us are unsure what that means in terms of how we should act. People spend time searching for goals, and there are institutions, most notably organized religion, which try to convince people that certain goals are more desirable than others.

In the early part of this century, Frank Knight emphasized the instability of goals. He wrote:

> "Wants...not only *are* unstable, changeable in response to all sorts of influences, but it is their essential nature to change and grow; it is an inherent inner necessity in them. The chief thing which the commonsense individual actually wants is not satisfaction for the wants which he has, but more, and *better* wants."[1]

Knight argued that "[e]conomic activity is *at the same time* a means of want-satisfaction, an agency for want- and character-formation, a field of creative self-expression, and a competitive sport."[2]

These problems of defining goals have played little role in the way economists have gone about their business. Milton Friedman, one of Knight's students, states the way most economists proceed:

> "Despite these qualifications, economic theory proceeds largely to take wants as fixed. This is primarily a case of division of labor. The economist has little to say about the formation of wants; this is the province of the psychologist. The economist's task is to trace the consequences of any given set of wants. The legitimacy of and justification for this abstraction must rest ultimately, as with any other abstraction, on the light that is shed and the power to predict that is yielded by this abstraction."[3]

---

[1] *The Ethics of Competition* (Univ. of Chicago Press: 1976) p. 22.

[2] *Ibid.* p. 42.

[3] *Price Theory: A Provisional Text* (Chicago: Aldine Publishing Co., 1962) p. 13.

People will consider as a **benefit** anything that moves them closer to the goal they are seeking. A businessman running a business will consider revenue as a benefit if his goal is to make a profit. A person striving for material advancement will consider more belongings a benefit. A father who wants his family to be happy will consider the joy of his child from a gift a benefit.

The discussion of goals has, up to this point, been qualitative, that is, not in terms of numbers. But economists have found that discussing goal-seeking in a quantitative way, that is, in terms of numbers, allows them to construct their theories much easier. As a result, economists talk about maximizing utility functions or objective functions. These are abstract ways, ways that are in mathematical terms, of saying that people are trying to attain goals. A utility function, written:

$$U = f(x1, x2, \ldots xn)$$

means that items x1, x2, etc. to some *n*th x all contribute to a person's utility. Utility, as the word is used here, is an abstract variable indicating goal attainment or want satisfaction. If a person has a simple goal such as accumulating material possessions, then x1 may represent cars, x2 fine furniture, x3 antiques, x4 land, etc. Anything which helps achieve goals gives, in the jargon of economists, utility.

Though it is of no practical importance here, it should be noted that many economists disagree with the interpretation given above. They see utility as a real psychic entity, just as happiness, joy, and satisfaction can be considered real psychic states. In this interpretation the possibility of measuring utility exists, though the techniques have not been developed. If, however, utility is a fiction invented to allow us to talk about goal attainment in a quantitative way, no general technique of measurement is possible.

Finally, economists have been reluctant to examine goals which involve competition for status. Perhaps their reluctance is due to their emphasis on the mutual advantages of exchange. In an honest transaction both the buyer and sellers must benefit or else the transaction will not take place. However, quest for status is zero-sum. If one person rises in status, he does so at the expense of others who he passes up. Their reluctance to examine goals involving status has meant that economists have surrendered some interesting questions to other disciplines. For example, to what extent does economic development depend on the goals which people have? Do some goals stop economic development? Limited evidence seems to suggest that groups in which it is socially unacceptable for a person to rise relative to his neighbors have a harder time developing than those groups in which social mobility is acceptable. Other social sciences have emphasized pursuit of status, and it gives them a very different way of seeing the world than the way economists do.

CONSTRAINTS AND COSTS

People have many goals, yet often those goals cannot be accomplished. This is because people face constraints or limitations on their behavior, and must maneuver within those constraints.

Constraints come in many forms. Sometimes they are mathematically imposed. Not everyone can be *above* average in intelligence, or athletic ability, or any other desirable trait. Not everyone can be a leader—there must be followers for leaders to exist. Not everyone can win, because the concept of winner implies that losers must also exist.

Time and biology impose the ultimate constraint on humans. A person's life is finite, almost always lasting less than a mere century. If one has many goals, or ambitious goals, the limited amount of time one has to live may make many impossible. This means, however, that the development of time-saving technologies, be it jet travel or packaged cake mixes, is liberating in the sense that it makes the constraint of time less pressing.

Income and wealth are other important

constraints of which people are aware. Often our goals require us to make other people act in certain ways. Lacking the ability to force them, one may try to persuade them, or one may pay them. If one wants bread from the baker or beer from the brewer, one pays them. Exchange of this sort takes place because both parties to the exchange find the exchange beneficial, i.e., it helps each side move toward its goals. Most people claim that their money and income is not sufficient and that they need more, which means that they would like to influence the actions of many more people than they in fact can. Economists call this limitation the budget constraint, a concept you met in Chapter 2.

Budget constraints reflect more basic constraints. People have limited abilities and limited time in which to earn income. In order to earn income, people sacrifice leisure time. In general the existence of constraints means that people faced choices, and thus costs. It means that people cannot accomplish all their goals (satisfy all their wants), but must choose to forgo some goals in order to accomplish others.

This point is important enough to justify spending some time with examples. Consider the college student who on one hand wants to do well academically and on the other wants to have an interesting and exciting social life. The basic limitation that this student faces is time. Each day has but 24 hours, and each week has but seven days. If enough time is spent to achieve a really excellent grade-point average (GPA), say straight *As*, the student may have a poor or miserable social life, or a low fun-point average (FPA). If the student enjoys as high a FPA as time permits, there may be little or no time for study, and grades may be very poor. This notion of a **tradeoff**, that to get more of one thing, a person must sacrifice something else, is very important in the way economists view the world.

Figure 1 illustrates the tradeoff that the student faces. Points to the right of the line are not attainable, while those to the left of the line and points on the line are. Point *a* represents a use of time with a great deal of studying and very little social life. Point *b* represents the opposite. Point *c* represents a poor use of time—the student in this case may be trying to study when everyone else is socializing and trying to socialize when most others are studying. In this case a different use of time could improve both GPA and FPA.

The tradeoff line in Figure 1 is curved. This curve indicates that, starting from a position of no study and all fun, devoting only

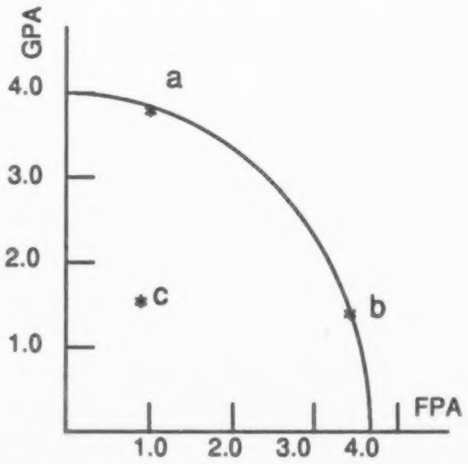


Figure 7.1 Higher Grades mean less fun

a few hours to study has a big effect on GPA but only reduces FPA only a little. Then for each additional hour spent studying the increase in GPA becomes less and less and the drop in FPA becomes greater and greater. In other words, if the student is not studying much, an extra hour with the books will help his GPA a lot and not cost him much in lost fun. If the student is studying a lot already, an extra hour with the books will not help his GPA much, but will cost a lot in lost fun. Economists call this pattern "decreasing returns" and find its presence in a great many situations.

In our case of GPA vs FPA, there is no money or prices. There are, however, costs. The cost of a higher GPA is the loss of FPA that must be given up. The cost of a higher FPA is the loss of GPA that must be given up. The notion of cost in economics refers not just to money costs, but to all options, whether measured in money or not, that must be sacrificed to get something.

Consider a different case, a case of books vs movies. Suppose that Rose Bole has a fixed amount of money to spend, say $100, and that Rose has two passions in life: she loves to buy books and to attend movies. If all books cost $5.00 and all movies cost $2.50 (these are simply assumptions to make the problem easier—as is the assumption that only two items are involved in the problem), then Figure 2 shows the options open to Rose. The budget line is a frontier showing what Rose can attain. The budget line limits choices; it is due to scarcity. The cost of a book is $5.00 or two movies. Spending money on a product means that money cannot be used to purchase another product. In the case of books vs movies, the tradeoff is a straight line because one more book always costs two movies, regardless of how many books Rose has already.

You should be able to see that the slope of the budget line depends only on the price of books relative to the price of movies. If either books get cheaper or movies get more expensive, the budget line in Figure 2 will get steeper. If this is not immediately obvious, compute the possibilities open to a person with $100 to spend if books and movies both cost $5.00 (a case of more expensive movies), and the possibilities open to a person with $100 to spend if books and movies both cost $2.50 (a case of cheaper books). Graphing the possibilities open to a person with only

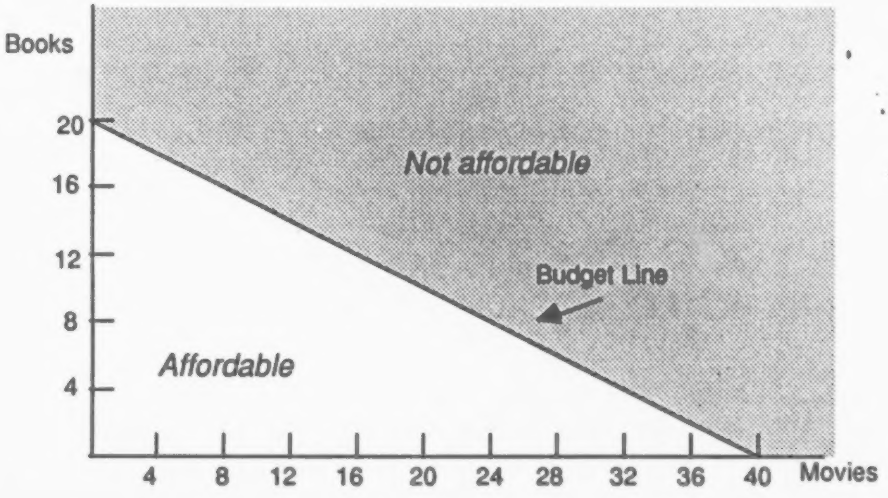
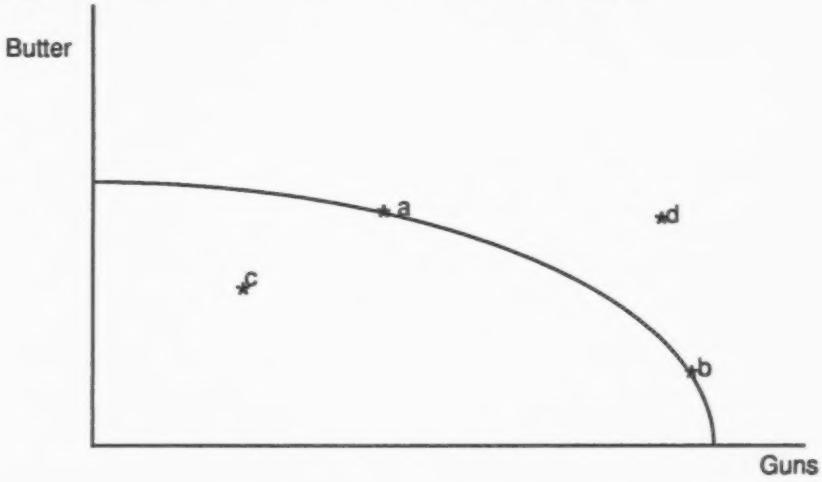


Figure 7.2 A typical budget line.

Figure 7.3 More for defense means fewer consumer goods.



$50 to spend but with books costing $5.00 and movies $2.50 gives you a line that is to the left of the line in Figure 2, but parallel to it, which means that it has the same slope. The amount of money available to spend does not determine the *slope* of the budget line; only the ratio of prices does that. (See question 3 at the end of the chapter.)

A final example of a constraint is that in the case of guns vs butter. During the Second World War the United States decided it needed to produce large amounts of armaments (guns). It shifted factories which had previously been producing goods for civilian use (butter) to the production of guns. This tradeoff could be represented as a move from a point such as *a* to a point such as *b* in Figure 3, except that at the start of the war there was still a high level of unemployment left over from the recessions of 1929-33 and 1937-8 (a period better known as the Great Depression). Hence the United States was not at the limit of what it could produce, but rather at a point such as *c*, which indicates that more of all goods could have been produced given the amount of resources and technology.

Though point *d* was a more desirable position than points *a* or *b*, it was unattainable given technology and resources. The limit to what is possible to produce is called the production-possibilities frontier, and is another concept you met in Chapter 2. Its existence, which is a result of scarcity, indicates that there are costs to producing all goods and services. During World War II the cost of producing thousands of tanks and jeeps was the virtual elimination of production of autos for civilian use. The cost of feeding millions of troops in the field was a less attractive diet for the civilian population.

The major idea in this section has been that all economic activity takes place within limitations or constraints. Because of these constraints, choosing results in sacrificed options. The options which are not taken, which sometimes can be measured in monetary terms, are costs.

THE MAXIMIZATION PRINCIPLE

A choice involves deciding in favor of one option and discarding others. The options that people have to choose from are limited by a budget constraint. To make the best decision a person must choose the option

TABLE 1 BENEFITS MEASURED IN UTILS

AMT.	UTILS FROM SHIRTS	UTILS FROM HAMBURGERS
1	11	8
2	20	15
3	27	21
4	31	26
5	32	30

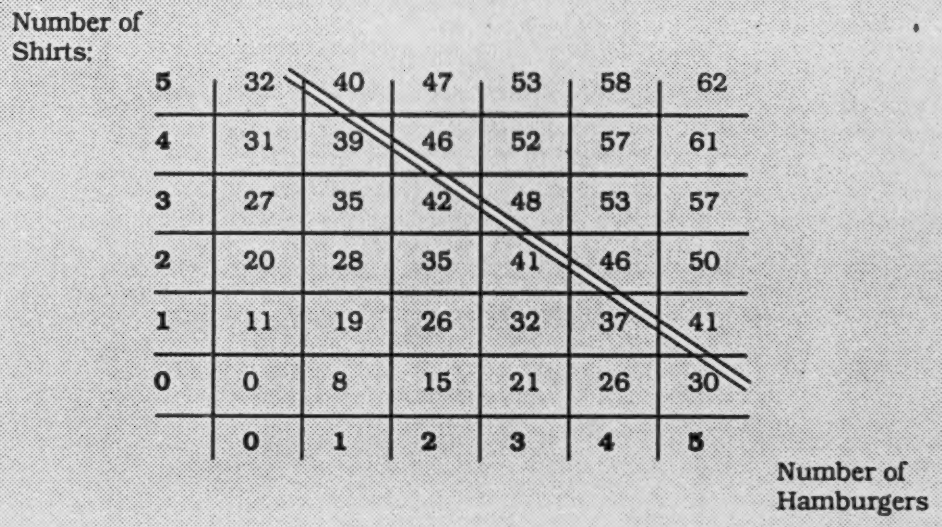
which is both possible and which contributes most to the achievement of that person's goals. This section analyzes how people can make such choices.

Though it is easy to show the budget constraint with a table or graph, showing goals is a bit more difficult. For the purpose of illustrating some important ideas, this section will assume that goal-attainment can be measured in some unit of satisfaction of utility. (The next chapter removes this assumption.) Table 1 gives an example using an imaginary measurement called the util.

Table 2 expands Table 1 to show utility for various combinations of shirts and hamburgers. Thus one shirt and three hamburgers give 32 utils of satisfaction (because 11 utils from shirts + 21 utils from hamburgers equals 32 utils). The person gets the same level of satisfaction from five shirts and no hamburgers. The person whose wants are described in this table should find these two combinations of equal value, or, to anticipate a term from the next chapter, he will be **indifferent** between them.

The consumer wants to get as much utility as possible, but a budget constraint limits him. In Table 2 the budget constraint is drawn so that the person can have only five items. Looking at all combinations possible, that is, to the left of the budget constraint, one can see that the combination three shirts and two hamburgers maximize utility. This combination yields 42 utils, and no other combination which is allowed by the budget

TABLE 2 A UTILITY FUNCTION



Number of Shirts:

5	32	40	47	53	58	62
4	31	39	46	52	57	61
3	27	35	42	48	53	57
2	20	28	35	41	46	50
1	11	19	26	32	37	41
0	0	8	15	21	26	30
	0	1	2	3	4	5

Number of Hamburgers

constraint gives more.

Often it is impossible or difficult to list all the options and the budget constraint as Table 2 does. There can be simpler ways to approach this problem. The beginning of this chapter suggested that we could break the question into a series of questions. We begin by assuming that all money is used to buy shirts, which in this example means that the person buys five shirts. Then we ask if the person is better off buying one hamburger than none. To answer this question, we need to compute the costs and benefits of making this change. The added benefits of the first hamburger is, according to Table 1, eight utils. To compute the cost of making this change, we must remember the budget constraint. To get a hamburger, the person must sacrifice a shirt. Since he began with five, he will be left with four. As a result, the utility he gets from shirts will decline by one util, and this is the cost of adding a hamburger. Since the change adds benefits of eight utils at a cost of one util, this is a smart change to make.

Now we ask if another change is worthwhile. What will the costs and benefits of adding a second hamburger be? Table 1 says that the utility of two hamburgers is 15. Since eight of those utils come from the first hamburger, the added utility of the second hamburger is seven. Since the budget constraint forces the person to give up his fourth shirt in order to obtain this hamburger, utility from shirts will drop from 31 utils to 27 utils, a loss of four utils. Thus, the benefits of adding the second hamburger are seven utils, and the cost is a loss of four utils. Adding the second hamburger is also a smart move because it increases total utility.

The third hamburger is not worth obtaining. The benefit of adding the third is six utils (moving from 15 to 21 utils in Table 1). But this move requires the person to move from three shirts to two, and in this move seven utils from shirts are given up. Since the cost of adding the third hamburger (seven utils) are greater than the benefits of this hamburger (six utils), the person should not add it.

Economists call the approach taken in the preceding paragraphs the **marginal** approach. Thinking on the margin means that a person is asking what the effects of small changes will be. In this approach one considers **marginal costs** and **marginal benefits**. The marginal cost of a change is the change in costs caused by the change. The marginal benefit of the change is the change in benefits caused by the change. The marginal approach suggests that one should make all the changes which increase benefits more than they increase costs (or which reduce costs by more than they reduce benefits). When all these changes have been made, one will find oneself at a point for which marginal costs equal marginal benefits. This rule for finding the best level of an activity is called the **maximization principle**.

To see that the maximization principle

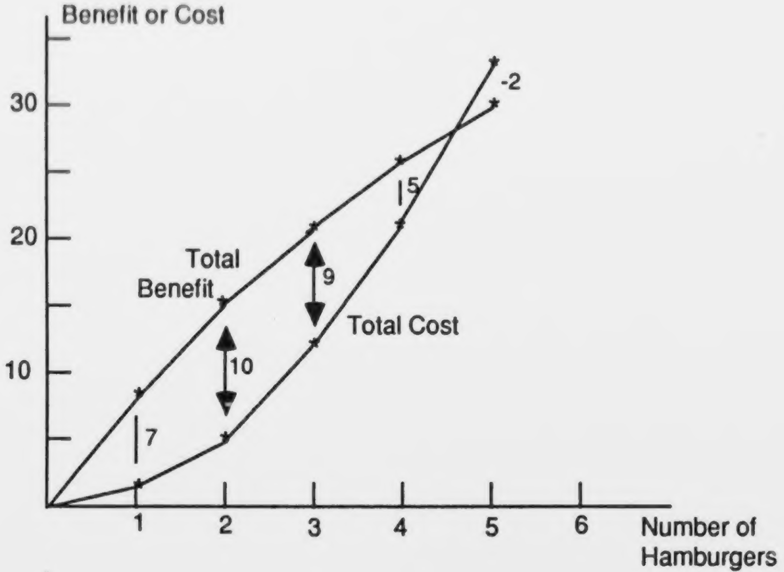
TABLE 3 COSTS AND BENEFITS OF HAMBURGERS

Number of Hamburgers	Marginal Benefit of Hamburgers	Marginal Cost of Hamburgers	Total Benefit of Hamburgers	Total Cost of Hamburgers	Net Benefit
1	8	1	8	1	7
2	7	4	15	5	10
3	6	7	21	12	9
4	5	9	26	21	5
5	4	11	30	32	-2

does generate the largest net benefits, the problem of how many hamburgers to buy can be analyzed with total costs and total benefits. This analysis is illustrated in Table 3. Columns two and three show marginal costs and benefits, and the way in which they were obtained has been described in the previous paragraphs. Total benefits of hamburgers are taken from Table 1.

Total Costs are obtained from column two of Table 1 and depend on the budget constraint. The total cost of three hamburgers, for example, will be the lost utility of three shirts. Since five shirts give 32 utils, and losing three leaves only two giving 20 utils, the total cost of three hamburgers is 12 utils. The net benefit column in Table 3 is found by subtracting total cost from total benefit. At

Figure 7.4 Total cost and total benefit curves.



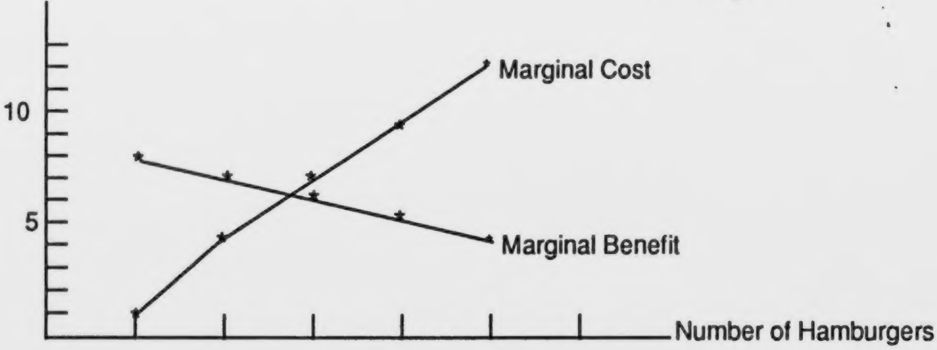


Figure 7.5 Marginal cost and marginal benefit curves.

two hamburgers total utility will be ten utils higher than at the starting point of five shirts and no hamburgers.

It is easy to see that if one has total cost one can obtain marginal cost, and if one has total benefit one can obtain marginal benefit, and vise versa. The formula for marginal cost is:

Marginal Cost = (Change in total cost)/
(change in activity)

Thus if a business knows that the total cost of producing 98 shirts is $398 and the total cost of producing 100 is $400, the marginal cost of the 100th shirt is approximately $2/2 = $1.00. Notice that marginal cost is *not* the same as average cost, which is found by dividing total cost by output. Alternatively, if one knows marginal cost or benefit, one can find total cost or total benefit by adding up all the marginals. (Check Table 3 to see that this is so.)

These results can also be shown graphically. In Figure 4 the total costs and benefits from Table 3 have been graphed. The goal of the person, to maximize net benefits, requires that the person try to find the point where the total benefit curve is at its greatest vertical distance above the total cost curve. (Here is one case in which the person does *not* want to end up at the intersection. Can you see why?) At this point, the total cost and total benefit curves have the same slopes. Before this point the total benefit curve is steeper, so they are moving apart. After this point they are moving together, which means that the total cost curve has the steeper slope.

The slope of the total benefit (or cost) curve is the rise over the run, or the change in total benefit (or cost) divided by the change in the number of hamburgers. But the marginal benefit (cost) of hamburgers is also defined as the change in total benefit (cost) divided by the change in the number of hamburgers. Hence, the slope of the total benefit curve is marginal benefit and the slope of the total cost curve is marginal cost.

This idea is used to construct the marginal benefit and marginal cost curves in Figure 5. The marginal curves are obtained by graphing the slopes of the total curves. The point at which they cross corresponds to the level of activity for which the slopes of the total cost and total benefit curves are equal.

THE EQUIMARGINAL PRINCIPLE

At this point you may think we have exhausted all the insights we can get from the hamburger-shirt problem in Table 1. Alas, there is still one more. Table 4 contains columns showing the marginal utility of shirts and the marginal utility of hamburgers (ignore the other columns for now). These marginal utilities are obtained from Table 1, which shows the total utility of one shirt, two shirts, etc. Marginal utility is the utility of the *first* shirt, the *second* shirt, etc. Thus, the utility of the fourth hamburger is found by subtracting the utility of four hamburgers from the utility of three hamburgers. Notice that the marginal utility of each good declines as more of it is used. This is a case of diminishing returns that has the special title of "the law of diminishing marginal utility." It is based on everyday observation and introspection. After four beers, a fifth gives less pleasure than the fourth, a third hamburger gives less satisfaction than the second, etc.

Suppose that the person is not at the optimal solution of three shirts and two hamburgers. Suppose instead that he has two shirts and three hamburgers. Can we tell from the table that he has spent his money incorrectly?

We can. Shirts and hamburgers cost the same. Suppose each costs $1.00 and the person has $5.00 to spend. Then the last dollar spent on hamburgers gave the person only five utils, while the last dollar spent on shirts gave him nine utils. The dollar spent on shirts gave a much larger return, and if he could shift money from the area in which it is giving a low return to the area in which it has a high return, the person will be better off. This is the basic idea of the **equimarginal**

principle. Maximization occurs when the return on the last dollar spent is the same in all areas. In terms of a formula, a person wants:

$$\frac{\text{Marg. Benefit of A}}{\text{Price of A}} = \frac{\text{Marg. Benefit of B}}{\text{Price of B}}$$

The power of this idea can be shown if we change the original problem. Suppose that the person still has $5.00 to spend, but the price of shirts doubles from $1.00 to $2.00. The old solution of three shirts and two hamburgers will no longer be affordable but will lie to the right of the budget line. To solve this new problem, two new columns must be added to Table 4, the marginal utility of shirts per dollar and the marginal utility of hamburgers per dollar. Table 4 does this in columns MUs/Price (the marginal utility of shirts divided by the price of shirts) and MUh/Price.

The equimarginal principle tells us to maximize utility by selecting, until our budget is used up, the highest values in the columns giving marginal utility per dollar. A person with only two dollars should buy two hamburgers rather than one shirt because both eight and seven are larger than five and one half. A person with $5.00, as in our example, should buy three hamburgers and one shirt. This decision does not quite equalize returns on the last dollars spent on shirts and hamburgers, but it comes as close as possible. Any other combination would give less utility and would allow for further improvement. For example, if one bought two shirts and one hamburger, the extra satisfaction from a dollar spent on shirts is only four and one half utils, while shifting money to hamburgers would allow one to get seven utils per dollar.

The equimarginal and maximization rules are closely related, but not quite identical. The next chapter will explore their relationship.

THE RATIONALLY IGNORANT VOTER

Marginalism is an application of the basic idea of calculus, and though calculus was invented a century before Adam Smith, it was a century after Smith when economists realized its significance. This "marginalist revolution" greatly clarified economic theory. The better understanding of their theory has prompted economists to search for new areas in which to use it. We conclude this chapter with visits to two areas economists have explored fairly recently.

TABLE 4
THE EQUIMARGINAL PRINCIPLE, OR HOW TO SPEND YOUR LAST DOLLAR

NUMBER	MARGINAL UTILITY OF SHIRTS	$\frac{MU_S}{PRICE_S}$	MARGINAL UTILITY OF HAMBURGERS	$\frac{MU_H}{PRICE_H}$
1	11	$5\frac{1}{2}$	8	8
2	9	$4\frac{1}{2}$	7	7
3	7	$3\frac{1}{2}$	6	6
4	4	2	5	5
5	1	$\frac{1}{2}$	4	4

When a person makes a major purchase, say of a new car, the person usually spends time learning about various makes of car and shopping for prices. Efforts spent in these activities can be seen as a way of producing knowledge. The more time spent, the more knowledge gained. Time, however, is limited, and spending time searching for information means that one cannot use that time for other purposes.

There is a limit on how much knowledge is worthwhile to gain. After some amount of reading, talking to friends and acquaintances, and visiting automobile dealers, a person finds that the extra benefit of another hour spent on these activities is less than the value of that hour. When one judges that this point has been reached, one stops searching and makes a decision.

The amount of time people spend obtaining information differs from product to product. They will spend less learning about the bicycle they give their child than they will learning about a new car, less in deciding which brand of soup to buy than in deciding which house to purchase, and less in deciding which brand of dog food is best for Rover than in finding a college for their first-born. The larger the purchase, the larger the potential benefit of a few hours spent learning about the purchase.

The government has many policies that involve major sums of money. For example, a major weapons system in the defense department can cost $50 billion dollars. This amounts to about $200 for every person in the United States, or $1000 for a family of five. Yet few people spend much time studying these policies. A reason is that to understand them requires many hours of study, and the probability that an understanding of them will change them in any way is very small. Thus, for most citizens the benefit of learning about a program that does not directly affect them is small, the cost is large, and as a result they do not know much about the program. Economists say that these poorly informed citizens are **rationally ignorant**.

The phenomenon of rational ignorance is not confined to political affairs. There is vastly more to know than any one person can possibly know. To survive and prosper in the world, one must seek that knowledge which will benefit the seeker. Most people would consider a bit odd someone who was not planning to buy a car but went from dealer to dealer trying to learn all he could about relative car prices in the name of intellectual curiosity. The behavior of most citizens suggests that they also consider odd the seeking of in-depth knowledge about the pros and cons of a specific government policy if that knowledge does not directly benefit the person who gets it. The hypothesis of the rationally ignorant voter suggests that people will be better informed about the choices they make in the marketplace than about those they make in the voting booth.

A look at costs and benefits not only explains why few citizens understand the subtleties of most government policies, but it also explains why about one half of the eligible voters in the United States do not vote. The probability that one's vote will be the crucial vote that decides an important election is small. Even if one's vote is the crucial vote which breaks the tie, one may not like the outcome—many people regret the way they voted when they compare actual performance with campaign promises. Given these small benefits compared to the costs of time and transportation that voting entails, it is not surprising that many people who are eligible to vote do not. What is surprising is that the percentage of people voting is not even smaller. It seems likely that there are other benefits to voting that have not yet been mentioned.

Politics is in many ways a spectator sport, with all the excitement and drama of football or baseball. Voting may be enjoyable in the same way as watching and cheering on a favorite ball team. Indeed, voting against a politician one does not like is enjoyable even if it does not result in his defeat. Another explanation for voting is that people have a sense of public duty. They want to be good citizens, and voting may seem important regardless of its effect—the act of voting itself can be important as a symbolic act. One other possibility is that people may overestimate the importance of their vote and the probability that theirs will be the ballot that decides an election.

In contrast to elections in the United States, elections in Eastern Europe are predictable. There is no doubt about who will win. Yet these countries report that impressive percentages, sometimes over 99%, of their citizens vote. A student who understands how to reason in terms of costs and benefits should be able to explain the implications of very high participation rates in meaningless elections.

Keep the rationally ignorant voter in mind when interpreting polls that ask citizens their opinions about complex public issues. The idea that voters are rationally ignorant also has implications for how governments work, the topic of Chapter 13.

SLAVERY

Slavery has intrigued economists who use economic theories to try to understand the past. These economic historians have painted a picture of slavery which differs considerably from the conventional historical account of the 1950s and 1960s. A book by Robert Fogel and Stanley Engerman called *Time on The Cross* (Little, Brown and Co., 1974) has summarized much of this research.

A main thesis of Fogel and Engerman is that slavery persisted because it was profitable for the slave owners. The conventional wisdom prior to studies of economic historians (or cliometricians—those who use statistical methods to study history—Clio was the muse of history in Greek mythology) was that slavery was not profitable, and that either it would have died away without the Civil War, or that it persisted because owning slaves carried benefits of honor, prestige, and power. On this point the evidence of the economic historians which Fogel and Engerman summarize seems to have carried the day.

One of the most controversial parts of Fogel and Engerman's book dealt with the question of how well slaves were treated. Fogel and Engerman maintain that slave owners used a carrot-and-stick approach to maintaining the productivity of slaves, and that many previous writers neglected the importance of the carrot part of the approach. Slaves who performed well were given prizes, holidays, cash, and bonuses. They were sometimes given land to cultivate for their own use, or they could make items for sale and keep the proceeds. Hard work increased the chance that a slave would be trained as an artisan (blacksmith, carpenter, etc.) or become a driver or overseer. (Many overseers were black.) Slaves who did not perform well, on the other hand, were denied these privileges and saw the stick part of the approach as they were confined to stocks or a jail, branded, sold, or whipped.

Fogel and Engerman argue that slave owners tried to balance rewards and punishments. Beyond some point additional punishment would reduce production, as slaves would become sullen and discontent. Rewards, on the other hand, could not be too generous, or else they would cost more than the gain from improved performance. The masters, according to Fogel and Engerman, were engaged in a maximizing behavior, trying to set the marginal benefits from rewards and punishments equal to their marginal costs.

Fogel and Engerman assert that the punishments used can also be explained in terms of maximizing behavior. Slaves were rarely punished by withholding food, because this would harm both slave and owner. Nor was execution a common punishment, even for murder. Since executing a slave resulted in a loss for the master, a more common treatment was a severe whipping and sale of the slave in another state. The cost of whipping, on the other hand, was borne entirely or almost entirely by the slave. Whipping died out as a punishment in the North with the development of the factory system because a factory owner had a different way of punishing. He could fire an offending worker, an option that made no sense to a slave owner. In a free labor market any inconvenience which the fired worker experienced, whether loss of food, clothing, or shelter, did not affect the factory owner in any way.

The picture Fogel and Engerman draw of slavery is that of slave owners trying to maximize their profits. They attempted to structure rewards and punishments so that the economic performance of the slaves was maximized. They were apparently successful, because slaves were more productive than free men in agriculture. Though this may seem to be a shocking conclusion, contradicting what many history books teach, it does have some common-sense appeal to it. Slave owners had a wider variety of incentives available to them than did an employer of free men. Slave owners not only could reward good behavior, but they had the option of punishing poor behavior. This option to use *force* was not available to an employer of free labor.

Fogel and Engerman contend that the positive incentives used to obtain the high productivity of slaves resulted in a material standard of living for slaves which was quite good compared to that of free workers. Thus it would seem that a system which is widely regarded as immoral performed well economically. This implication seems to be at the heart of some of the controversy which has surrounded *Time on The Cross*. The world would be a more comfortable place if we could be sure that bad people and bad systems lose or fail. Yet Fogel and Engerman do not attempt to justify slavery; in fact, they find it economically inefficient. Chapter 11 will explain the paradox of how a system can simultaneously be very productive and very inefficient.

SUMMARY

This chapter considered the mathematical approach to making decisions on the basis of weighing costs and benefits. If one knows the costs and benefits of an activity, there are rules which will tell one the level of the activity to choose so as to get the greatest net benefit. This chapter introduced you to these rules in the form of the maximization principle and the equimarginal principle.

Though the rules which this chapter discussed can be derived mathematically and are in fact nothing but applied calculus, they also make common sense. Since you will see them frequently in the upcoming chapters, it is important that you understand the logic and common sense behind these rules.

QUESTIONS

1. Consider the graph in Figure 1 and the accompanying text which illustrates the tradeoff in the case of GPA vs FPA. How could you illustrate on this graph the difference between two students who were identical except that one was more intelligent than the other? How could you illustrate the difference on this graph between two students who were identical except that one had to work 20 hours a week in order to pay college expenses?

2. Chapter 1 introduced you to the concepts of positive and normative statements. Which of the following are positive statements, and which are normative?
   a. A system of slavery produces as much as a system of free labor.
   b. Because it is productive, a system of slavery is good.
   c. Slavery is a bad system because it is unfair.
   d. A good citizen will keep well informed on political issues.
   e. Everyone should vote.
   f. Many people do not find it in their interests to vote.

3. The graph in Figure 2 shows a budget line for books and movies assuming $100 to spend, the price of books equal to $5.00 and the price of movies equal to $2.50. Re-draw this graph, then show how it will change if (a) the price of books is cut in half to $2.50; (b) only $50 is available to spend; (c) the price of movies doubles to $5.00.

4. Joe likes Tom Thumb's Tiny Burgers. For him they satisfy an important want. The schedule below shows how much "want-satisfying" power they have:

# of Burgers	Want-Satisfying Power
0	0
1	5
2	9
3	12
4	14
5	15
6	15
7	14

a) If Tom Thumb's Tiny Burgers are free, how many should Joe take?

b) Suppose that for every Burger Joe buys, he must give up something that has 2.5 units of want-satisfying power. How many should he buy?

5. Hans Fruehstueck has $1.00 to spend on breakfast. He must decide how many donuts and cups of coffee to buy. The utility these items give him is shown in the table below:

Number	Utility of donuts	Utility of coffee
1	5	10
2	10	13
3	14	15
4	16	16

a) Construct a table similar to Table 2 and decide what Hans should do if coffee and donuts both cost $.25.

b) If coffee and donuts both cost $.25, construct a table similar to Table 3 and find how many of donuts he will buy by computing the marginal cost and marginal benefit of buying donuts.

c) Construct a table similar to Table 4 and show how it tells you what Hans should do.

d) How will you find Hans' new best position using the methods of parts a, b, and c if the price of donuts rises to $.50? If he only has $.75 to spend?

6. Suppose a person has $10.00 to spend, the price of apples is $2.00 and the price of bananas is $2.00. The utilities from apples and bananas are shown in the table below. a) Based on these data, how many apples and bananas should the person buy? (Hint: You will need to create three new columns to answer this. Look again at Table 4, and recall how one gets marginal utility from total utility.) b) If the price of apples drops to $1.00, how many of each will the person buy? (Hint: you must add still another column.) c) If the price of apples rises to $4.00, how many will the person buy? d) Using these answers, draw a demand curve for apples. e) What would an increase in income do to this demand

curve? f) What would a change in tastes do the demand curve and to the table below?

# of Apples or Bananas	Marginal Utility of Apples	Total Utility of Bananas
1	20	30
2	19	55
3	18	76
4	17	94
5	16	110
6	15	124
7	14	136
8	13	146
9	12	154
10	11	161

7. Using the data below, compute the level of output for which net benefits are at a maximum.

OUTPUT	AVERAGE BENEFIT	MARGINAL COST
1	20	$6\frac{1}{2}$
2	18	7
3	17	$7\frac{1}{2}$
4	15	8
5	13	$8\frac{1}{2}$
6	11	9

Explain how you got your answer. (Hint: average benefits are total benefits divided by output.)

8. What do the very high election turnouts in Eastern Europe imply?

9. Answer the questions below using these data:

Amount	MUa	MUb	MUc
1	26	20	34
2	18	16	30
3	16	12	8
4	14	8	2
5	12	0	0

Price of A = $2.00
Price of B = $2.00
Price of C = $4.00

a. According to the above table, what is the total utility of consuming 3 units of item C?

b. The person with these marginal utilities will not buy 4 units of A unless he has at least how much to spend?

c. If a person with the marginal utilities above has $16.00 to spend, how much of each should he buy to maximize utility?

d. If the price of A rises to $4.00 and the person still has $16.00 to spend, what happens to the optimal amount of C?

CHAPTER 8

MAXIMIZING BEHAVIOR

After you read this chapter and work the problems at the end, you should be able to:

Interpret information on a graph showing a budget constraint and indifference curves.

Define consumers' surplus and explain how it is measured on a demand curve.

Define producers' surplus and explain how it is measured on a supply curve.

Explain the paradox of value.

Compute the present value of a future sum when given an interest rate.

Explain what happens to the price of bonds when the interest changes.

This chapter applies and extends the analysis of the previous chapter about maximizing behavior. It begins by abandoning the assumption of the last chapter, that utility can be measured. It then examines some ideas that are extensions of the logic of maximizing behavior: consumers' and producers' surplus, in-kind gifts, and present value.

INDIFFERENCE CURVES AND DEMAND

Economists write the notion that using goods and services helps people achieve goals (attain utility) as a mathematical function, or

$U = f[x1, x2, . . . xn)$

Table 2 in Chapter 7 shows one way to represent this equation. It gives a "map" of consumer preferences. However, there are two problems with this table. First, it assumes an independence between goods that seems unlikely. The added utility of a shirt depends only on the number of previous shirts that have been bought, and not at all on the number of hamburgers. If shirts and hamburgers are either substitutes or complements, this independence should not exist. However, a table similar to Table 2 can easily be made that will avoid this problem.

The second problem is more serious and has led economists to avoid tables (except for explanatory purposes) when discussing utility theory. A table such as Table 2 assumes that utility can be measured. Since utility refers to the extent that goals are accomplished, there is no way of knowing whether any measurement is arbitrary or not. This problem can be overcome by using graphs or mathematics.

At first glance, using a graph instead of a table may not seem like a good way to proceed. Discussion of maximizing utility must involve at least three variables: the amount of good *A*, the amount of good *B*, and the level of utility. Graphs have only two axes, and three variables seem to require a three dimensional construction rather than a two dimensional one. However, there is a way around this problem, one that geographers use when they draw contour maps. They also

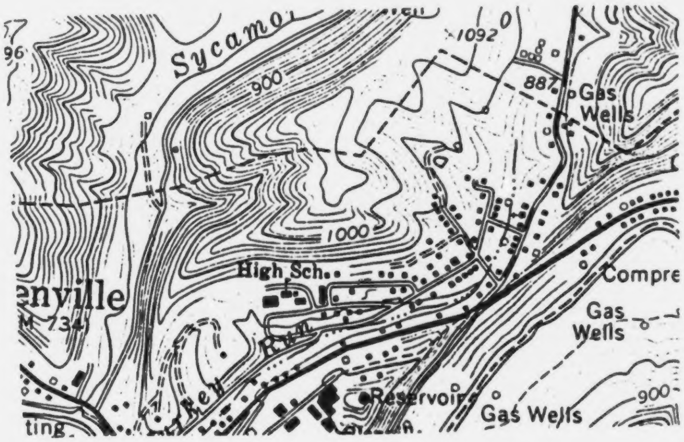
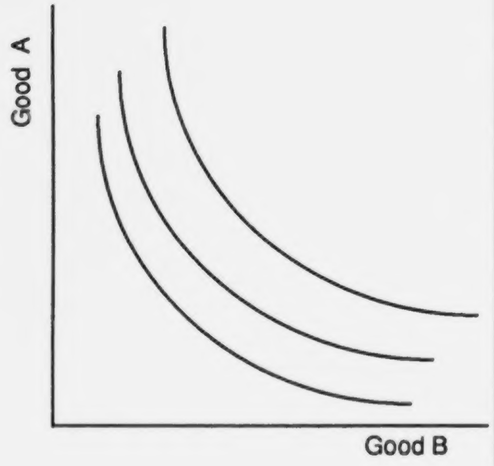


Figure 8.1 Contour lines showing altitude.

Figure 8.2 Isoutilility or indifference curves



must show three variables: longitude, latitude, and altitude. They show altitude with a series of lines or topographic contours such as those in Figure 1 which shows a hilly section of West Virginia.

The same method of construction can be used to show a utility map. A line will connect all possible combinations of good *A* and good *B* which show the same level of utility. This line is called an **isoutility** (*iso*-is Greek and means "the same" or "equal") line or an **indifference** curve. In general these isoutility lines will have a curvature as they do in Figure 2 if diminishing returns holds. Notice that one does not need to measure utility in order to draw a graph such as that in Figure 2. All one needs is the ability to *order* different levels of utility, that is, to say that bundle *A* is preferred to bundle *B*, or that bundle *B* is preferred to bundle *A*, or that the chooser is indifferent between the two bundles.

Indifference curves assume that individuals are consistent. If Jane prefers option A to option B, and if she also prefers option B to option C, then she should prefer option A to option C. Inconsistent behavior is not compatible with attaining goals. If people do not try to attain goals in a consistent manner, one can expect few regularities in their behavior.

To show what the consumer should do to maximize utility, a budget line must be added to the preferences shown in the indifference curves. Figure 3 adds one. Point *a* is not attainable because it lies to the right of the budget line. The consumer is indifferent between points *b* and *d* because they lie on the

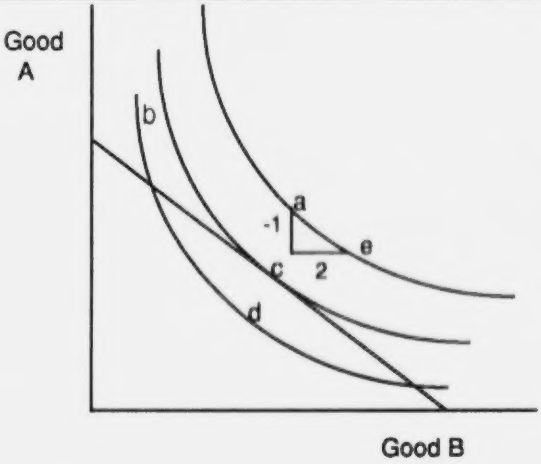


Figure 8.3 Indifference curves with a budget constraint.

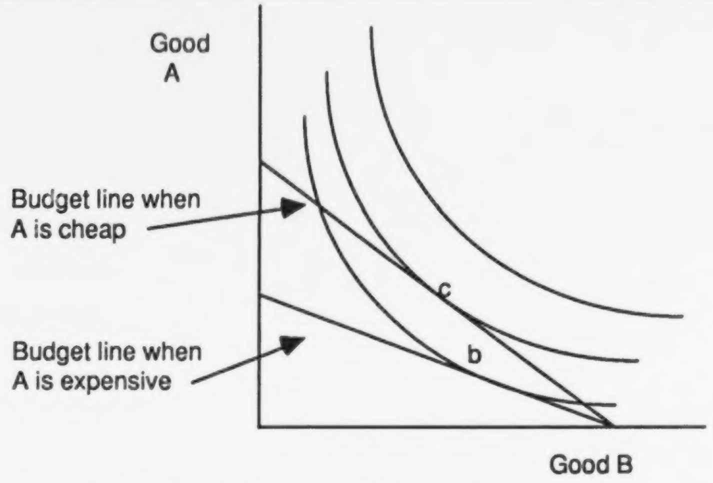


Figure 8.4 Effects of shifting the budget constraint.

same indifference curve, but point *d* is cheaper than *b* because *d* lies below the budget line. The consumer wants to get on the highest indifference curve affordable, and this will lead him to point *c*.

The effect of a rise in the price of good *A* is shown in Figure 4. A higher price of *A* means that less of *A* can be purchased, and hence the budget line moves to the left, intersecting the vertical axis at a lower point. Point *c* is no longer possible, and the consumer must move to a new position, which assuming utility maximization, will be point *b*. Unless the indifference curves are peculiar, point *b* will represent less of good *A* than will point *a*, which is what the law of demand says will happen.

Looking at two different prices has produced two different points on an individual's demand curve. By varying the price of good *A*, other points could be found and an entire demand curve for one individual consumer constructed. The market demand curve of Chapter 4 is obtained by adding up demand curves of all individuals.

The theory of consumer choice which the indifference curves embody is an elegant construction with which economists frame problems. It is an example of a nonscientific theory—it begins with a set of assumptions about how people act, but there are no outcomes which it prohibits (other than those which contradict the original assumptions). Though a downward sloping demand curve can be derived from it, so too can an upward sloping demand curve. Further, in recent years there has been a realization among economists that a picture such as Figure 3 may not be a good description of the decision-making process when people must make decisions with partial information, with fuzzy goals, under conditions of risk and uncertainty, and when options are difficult to compare. (See question 8 at the end of the chapter.)

IN-KIND AND CASH TRANSFERS

Suppose an eccentric millionaire decides to help a poor neighbor by giving him $1000 worth of nontransferable hat certificates. (The certificates could be used to buy hats, but nothing else.) Is this gift of the same value to the poor neighbor as a gift of $1000 in cash? Economists answer that generally a gift **in kind** has less value than a cash gift because it has restrictions. A cash gift gives more options, and economists usually assume that more options never harm a person, but may help one.

The argument that cash gifts are superior to in-kind gifts can be shown with budget lines. In Figure 5 the poor neighbor originally faces budget line A-B. A gift of $1000 worth of hat certificates means the person could buy more hats than he could previously have bought, or if all income were spent on hats, he could buy O-E rather than the O-B that he could have had before. There is no increase in Other Goods that he can buy if he spends all his income on them. As a result, the new budget line is A-C-E. A gift of $1000 in cash, on the other hand, would increase not only the number of hats that is possible, but also the amount of Other Goods. The new budget line with a cash gift is D-C-E. The dotted portion D-C represents the options that cash gives that the in-kind transfer does not allow. Since few people spend very much of their incomes on hats, most people prefer to be on the segment D-C rather than C-E. In this case the gift in kind is less valuable to the recipient than a gift in cash. (See question 6 at the end of the chapter to finish this analysis.)

This analysis, as all economic analysis, has underlying assumptions, and if these assumptions do not hold, the above conclusion may not either. First, it assumes that there is no cost in making decisions. Making decisions often involves gathering information, weighing it, and worrying about whether the decision is correct or not. Anyone who has had to choose between two good job offers knows the agony that making a decision can entail. Further, some people recognize that they do not make good decisions and hire others to do so for them. This shortcoming explains the career of financial advisor, in which people draw up budgets for others and in some cases make the actual purchases. Thus, if the giver knows of items which the recipient would like but does not know about, or if the cost of making decisions is high, or if the individual is not capable of making decisions that get him to his goal, the recipient may be better off with the in-kind transfer than with a cash transfer.

Most gifts between friends are not for cash but are in-kind gifts. Most girls would be

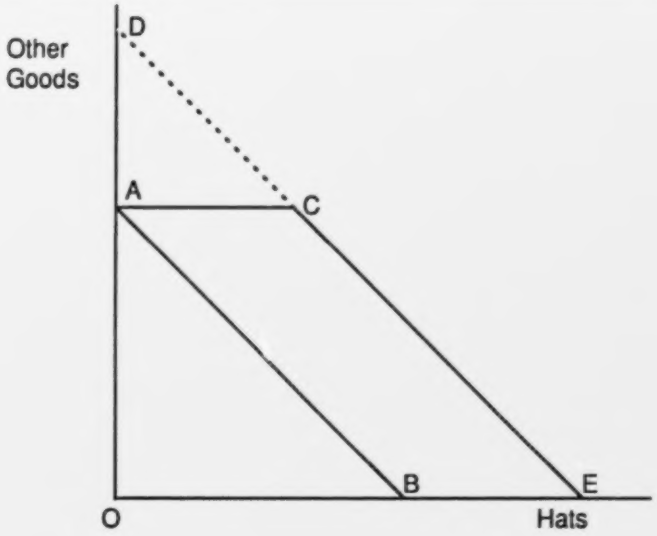


Figure 8.5 Budget constraints for cash gifts and in-kind gifts.

upset if their boyfriends gave them money for birthdays or Christmas. Gifts between friends are a way of cementing ties and give rise to obligations. They play an important role in helping small-group associations run smoothly. An in-kind gift indicates that the giver is interested enough in the recipient to learn about the recipient's likes and dislikes, and has spent time doing this and in obtaining the gift. The simple economic analysis of Figure 5 does not capture these subtleties. Though anthropologists have studied gift-giving more than economists, some economists do consider it important. For example, Kenneth Boulding's *Economics of Love and Fear* argues that there are three mechanisms that help hold groups together: the integrative mechanism of gift-giving and love, voluntary and mutually advantageous exchange, and coercion and threat.

Third, the economic analysis of Figure 5 refers only to the recipients preferences, not to those of the donors. A donor may disapprove of the goals that the recipient pursues, and thus may be unwilling to let him make choices. Economists are a bit unusual because they generally assume that people know what is best for themselves. As a result, most are reluctant to approve of desires of donors to deny choices to recipients.

The question of whether in-kind transfers are better than cash transfers is important when governments devote hundreds of billions of dollars to transfers. At one time the government gave food to the poor; now it gives food stamps. The in-kind vs cash argument was involved in the switch (though it may not have been the decisive factor). The government attempts to help poor people by building and providing public housing; would the poor be better off if the government provided money instead of the housing? The government provides education by providing free schools, but students can rarely choose which free school they attend. Would they be better off if the government stopped producing education and instead provided them with a "tuition voucher," a sum of money which they could use at the school of their choice? The present system of transfers is a mixture of cash and in-kind transfers that is only partially explainable in terms of the recipients' welfare.

PRESENT VALUE

Suppose that someone will give you a gift of $100, and will give this to you either now or in four years. Which is better, the money now or the money four years from now? The rule from the previous section, that usually gifts with restrictions are of less value than gifts without restrictions, suggests that money now is worth more than money in the future. Anything that one can do with the gift of $100 four years from now one can do with $100 now simply by saving it for four years. But there are many things that one can do with money now that one cannot do with money four years from now. Therefore, $100 promised four years from now is not worth $100 right now, but a smaller amount.

One of the things that can be done with money now is that it can be invested so that it will earn interest. Since this cannot be done with money four years from now, this option of foregone interest is a cost of waiting for the money. This cost can be measured and one can see the amount by which money in the future must be *discounted* to learn its present value.

Suppose that the interest rate is 10%. Then $100 can be turned into $110 one year from now. Thus, $100 now and $110 a year from now have the same value. (You may have to think about this for a while.) This simple idea is vital in business and governmental decisions because a great many decisions have costs and benefits spread over time, and it is often necessary to compare sums in different time periods.

Computing present value of future sums is nothing more than working compound interest problems backward. You should recall from your high school education that the formula for finding the future value of a present sum after one period is:

(1) $P + Pr = F$

or

(2) $P(1 + r) = F$

where P is the present sum, r is the interest rate in decimal form, and F is the future sum. (Try the formula for $P = \$100$ and $r = .10$. You should get $F = \$110$.)

After two years the amount of money will be:

(3) $F1(1 + r) = F2$

where F1 is the amount of money one year from now, and F2 is the amount of money two years from now. This may be rewritten as:

(4) $(P(1 + r))(1 + r) = P(1 + r)^2 = F2$.

(Try this formula for $P = \$100$ and $r = .10$. F2 should be \$121.) Using the same logic, the future value three years from now will be

(5) $P(1 + r)^3 = F3$

and for any arbitrary period *n* it will be

(6) $P(1 + r)^n = Fn$.

Simple algebra allows us to solve this equation if we have the time period and two of the three remaining variables. (Logarithms help a lot if we are solving for r.) In particular, if we have a future sum of money and want to find its present value, the last equation can be rewritten as:

(7) $P = \dfrac{F}{(1 + r)^n}$

Using this formula in our case of \$100 four years from now and an interest rate of 10%, the present value is:

$P = 100/(1.10)4 = 100/1.464 = 68.30$

This means that \$68.30 invested at 10% will grow to \$100 four years from now. \$68.30 now and \$100 four years in the future have equivalent value if the interest rate is 10%.

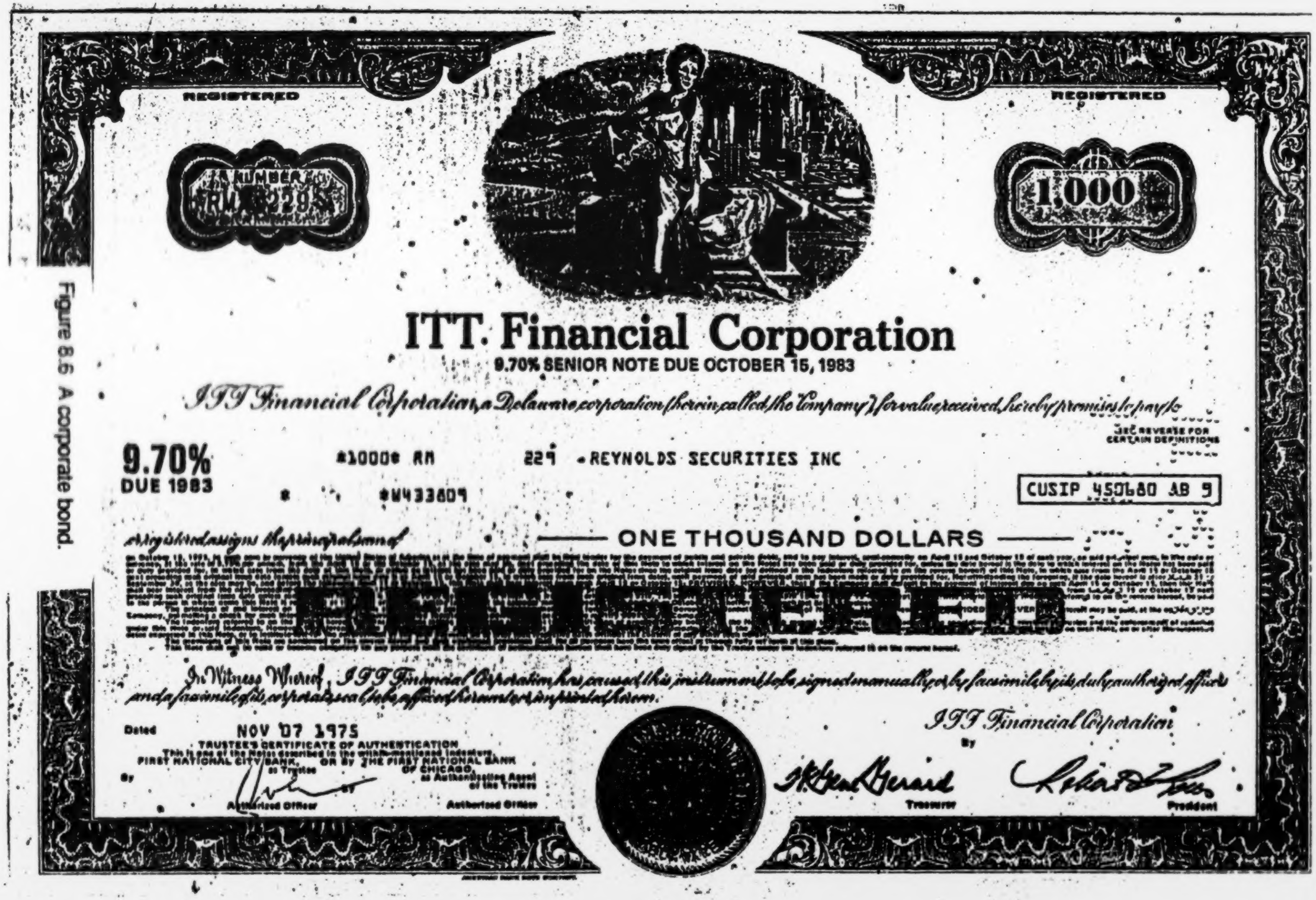
Present value explains why the price of bonds on the bond exchange falls when interest rates rise and rises when interest rates fall. A bond is a contract. At the beginning of the contract, the lender pays a specified amount, usually \$1000 or a multiple of \$1000, to the borrower. The bond specifies when the \$1000 will be paid back, and how much will be paid as interest each year and all other terms of the agreement. A bond issued in 1983, due in 2003, and paying \$130 a year interest (this means a coupon rate of 13%) has the stream of payments illustrated in Table 1. If the market rate of interest equals 13%, the 1983 value of the "get" column will equal \$1000.

TABLE 1
WHAT A BOND DOES

YEAR	GIVE	GET
1983	\$1000	\$0
1984	0	130
1985	0	130
1986	0	130
.	.	.
.	.	.
2001	0	130
2002	0	130
2003	0	130

If the lender decides that he no longer wants to hold this bond, he cannot demand payment from the borrower. The contract does not give the owner of the bond the right to payment on demand. But he can sell it to someone else, and the price of this sale need not be \$1000. If market interest rates have risen since the original purchase of the bond, then the present value of the future payments in the "get" column will drop, and so will the value of the bond. Another way of seeing this principle is to realize that if the market interest rate rises to 15%, there will be borrowers selling contracts for \$1000 that pay \$150 each year until the maturity of the bond (when the contract ends). It would be foolish to pay \$1000 to buy a contract that pays only \$130 a year.

The notion of present value may seem dry to someone who has never owned a bond or made business decisions. But for many cor-



REGISTERED

REGISTERED

NUMBER

RM 229S

1,000

ITT Financial Corporation

9.70% SENIOR NOTE DUE OCTOBER 15, 1983

ITT Financial Corporation, a Delaware corporation (herein called the "Company") for value received, hereby promises to pay to

SEE REVERSE FOR CERTAIN DEFINITIONS

9.70%
DUE 1983

1000 RM 229 •REYNOLDS SECURITIES INC

*W433609

CUSIP 450680 AB 9

or registered assigns the principal sum of —— ONE THOUSAND DOLLARS ——

REGISTERED

In Witness Whereof, ITT Financial Corporation has caused this instrument to be signed manually or by facsimile by its duly authorized officers and a facsimile of its corporate seal to be affixed hereunto or imprinted hereon.

Dated NOV 07 1975

TRUSTEE'S CERTIFICATE OF AUTHENTICATION
This is one of the Notes described in the within-mentioned Indenture.
FIRST NATIONAL CITY BANK, as Trustee OR BY THE FIRST NATIONAL BANK OF CHICAGO, as Authenticating Agent of the Trustee

By
Authorized Officer
Authorized Officer

ITT Financial Corporation
By
Treasurer
President

Figure 8.6 A corporate bond.

porate financiers, and for those who have money invested in bonds, it is a notion that provides a lot of excitement—both joy and woe—in their lives.

CONSUMERS' SURPLUS

The assumption that consumers maximize utility leads to the downward-sloping demand curve. Actually even nonrational or random behavior will lead to a downward-sloping demand curve, as economist Gary Becker has demonstrated, but this demand curve does not have the same interpretation that a demand curve based on utility maximization (trying to attain goals) has.

Becker's argument is quite simple. Because the budget line is a constraint separating what is possible from what is not possible, even nonrational consumers face a budget constraint. Becker notes that if people randomly purchase goods, they will be randomly distributed either along a budget constraint such as in Figure 2 of Chapter 7, or within the area bordered by the budget constraint. (Becker considers both cases.) If the price of a good increases, the budget line will shift, and a new random distribution of points will occur. The geometry of the situation implies that, on the average, people will buy less of a good as its price rises.

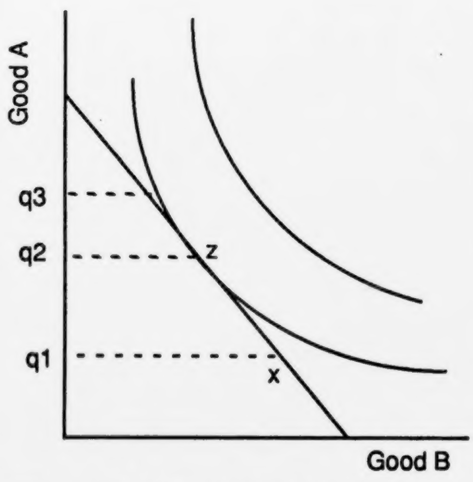


Figure 8.7 Utility Maximimization

Though the demand curve of nonrational consumers will slope downward, it can no longer be interpreted as a locus of points of consumer equilibrium. With the assumption of utility maximization, the preferences and prices used to construct Figure 7 imply that q2 is the amount of good A that is optimal for the consumer. If either more (q3) or less (q1) is being used, there is an incentive to change behavior because it would lead to better fulfillment of goals. However, if behavior is random and not concerned with fulfilling goals, point *x* is as good as point *z*. Thus the argument that price controls have unintended results depends on the assumption that behavior is goal-directed.

Utility maximization suggests that the demand curve, because it measures buyer's willingness to pay, shows marginal benefits to buyers. Table 2 indicates that people will buy only one item if the price is $5.00, or that people are willing to pay $5.00 for the first item. They are not willing to pay $5.00 for a second item, but only $4.00. A second item has a smaller marginal benefit than the first because of the law of diminishing marginal utility. The equimarginal principle suggests that as price gets lower, consumers find that they must use more of an item to keep equality among marginal-utility-to-price ratios. Or alternatively, as people use more of an item, its marginal utility drops, and so must its price if they are to stay in equilibrium.

This notion of the demand curve has an interesting implication known as the **consumers' surplus**. If in Table 2 consumers are buying three items, they must pay a total of $9.00. But the total value to them is $5.00 + $4.00 + $3.00 = $12.00. There is a surplus value of $3.00. In a more intuitive example, suppose a person has been working in the hot sun all afternoon and is extremely thirsty. This person may be willing to pay as much as $2.00 for a can of cold beer, but if he can buy

TABLE 2
A DEMAND CURVE

PRICE	AMT. PEOPLE ARE WILLING TO BUY
$5.00	1
$4.00	2
$3.00	3
$2.00	4
$1.00	5
$ .50	6

it for only $.50, he thinks he has found a good deal and may buy two or three. The difference between the maximum a person would pay and the actual amount that he does pay is consumer's surplus. In other words, consumers' surplus is the difference between the value in use of an item and its value in exchange.

Notice that consumers' surplus is not related to the type of surplus which occurs in a market when price is above market-clearing price. Perhaps economists would have avoided this possible confusion if they had used a term other than consumers' surplus for this concept, but they did not and the term is now well-established.

Consumers' surplus helps explain a puzzle, called the Paradox of Value, with which Adam Smith and other economists for a century after him struggled. This paradox notes that some items which have relatively little use to society, such as diamonds, are extremely expensive, while others which are vital, such as water, are cheap. Smith was unable to explain this situation adequately, and this paradox led to much theorizing that in retrospect was a detour.

To see how this paradox is resolved, consider again the downward sloping demand curve in Table 2. As an item grows more abundant, its total use value to consumers, which is the entire area under the demand curve, rises, but its price, or its marginal value to consumers, declines. Thus, if two items in Table 2 are available, total value to consumers is $9.00 (or $5.00 for the first and $4.00 for the second), but the price or value in exchange is only $4.00. If six are available, total use value rises to $15.00, but exchange value (price) drops to $.50. Smith and his early followers missed this distinction between marginal and total. Thus diamonds are scarce and have a high marginal value but a low total value. Another pound of diamonds has valuable uses that are not currently being met. Water is tremendously abundant and thus has a high total value and a low marginal value. Another gallon of water is not particularly important.

PRODUCERS' SURPLUS

People buy a product only when the price of the product is less than the value they put on the product. People sell a product only when they receive more than the value they put on the product. If there is a consumers' surplus, there should be a sellers' surplus. There is. It is called either **producers' surplus** or **economic rent**. The producers' surplus exists when actual price exceeds the minimum price sellers will accept. If a seller is willing to accept $2.00, but can get $5.00, the seller has a surplus of $3.00.

Producers' surplus can appear as profit, but usually it takes a different form. Suppose, for example, that the price of corn has been $2.00 per bushel for many years. Then it rises to $3.00 per bushel and stays there. This higher price will draw more land into corn production, but this change is of no importance here. What is of interest is what happens to the farmers who were producing corn at $2.00 per bushel, and now find that they can sell corn at $3.00. It certainly appears that these farmers are better off because a producers' surplus of $1.00 per bushel has appeared which was not there before.

However, let us separate farming into two parts: working the land and owning the land. Suppose a farmer does not own the land he works, but rents it. It then becomes unlikely that this farmer will benefit at all from the higher price of corn. If those working the land obtained the surplus, there would be competition for the right to work land that is especially suited to growing corn. This competition should raise the value of the land, and therefore it will be the landowners, not the cultivators, who benefit from the higher price of corn. Because the earliest case of producers' surplus analyzed was one in which land captured the surplus, the producers' surplus is often called economic rent.

Producers' surplus is usually captured by resource owners rather than by producers. Hence the producers' surplus is not the same as profit. The resources which capture the surplus are those which are especially good at producing the product in question or which have no other uses, and hence will be used for that product even when prices are low. Sometimes the resource which captures economic rent is labor. The high pay that superstars in many fields earn is mostly producers' surplus. The basketball star paid $1 million who would still play for $25,000 earns $975,000 in producer surplus. There is an interesting conclusion to this observation: the price of basketball tickets is not high because the salaries of the star athletes are so high (as owners sometimes allege), but the salaries are so high because fans are willing to pay so much for the tickets to see the stars play.

The producers' and consumers' surpluses are illustrated with supply and demand curves in Figure 8. The total value to consumers of quantity Q is represented by areas A+B+C. Since the consumers must pay B+C, only the area A is surplus for them. Producers get revenue of B+C. B is their surplus because only payments of C are needed to attract the resources necessary to produce quantity Q.

The concepts of consumers' and producers' surpluses are tools which can help analyze many situations. For example, is there any temptation for sellers to gang up on buyers? If sellers can raise the price, can they transfer some of the consumers' surplus to themselves? They can, and Figure 9 illustrates what happens. The consumers' surplus at price Pc is A+B+D. The producers surplus at this price is C+E. By raising price to Pm, sellers cause the consumers' surplus to shrink to the area A. Area B is transferred from consumers to producers. The new producers surplus is C+B. If sellers gang up on buyers, they are no longer price takers. Rather the sellers leave the supply curve and search along the demand curve for the best deal. As a result, such behavior is called "price searching."

It is easiest for sellers to restrict output and raise price when there are very few sellers and many buyers. When there is only one buyer, a **monopoly**, economists expect the seller to act in this way. With many sellers, coordination of decisions becomes difficult (for the same reason that the problem of the commons, discussed in Chapter 2, can exist) and output restrictions become unlikely.

Alternatively, buyers can gang up on sellers and extract producers' surplus. They must restrict purchases to drive the price down. Again, this behavior is likely only when there are very few buyers and many sellers. When there is only one buyer, a **monopsonist**, economists expect him to

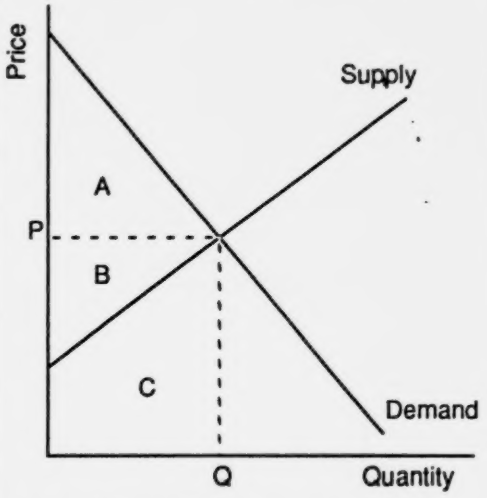


Figure 8.8 Producers' and consumers' surpluses.

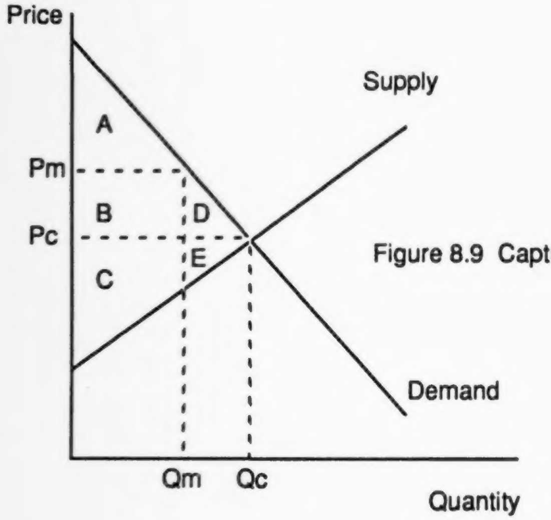


Figure 8.9 Capturing some of consumers' surplus.

restrict purchases.

Chapter 3 noted that what is good for the individual is not necessarily good for the group. Notice that the process of transferring the value of area B from consumers to producers in Figure 9 causes consumers to lose area D and producers to lose area E, and that no one gets this lost value. In the process of increasing their surplus by seizing area B, producers cause the value of total surplus to shrink. There is a conflict here between the interests of producers and consumers. But before we take a closer look at this conflict, we need to analyze the production unit of an economy, the firm. The next two chapters take up this task.

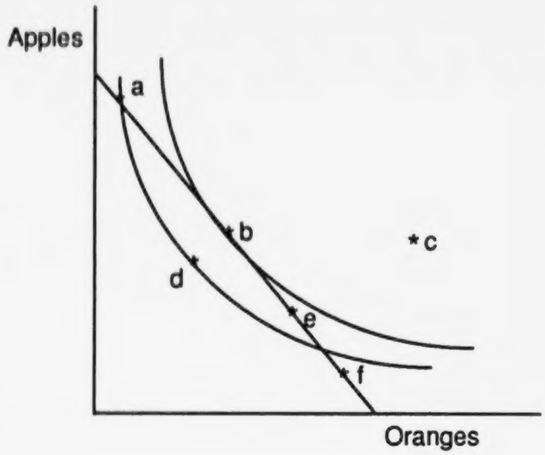
SUMMARY

This chapter explained how economists think about consumer choice. A set of indifference curves represent a person's preferences, and that person strives to find the point on the highest indifference curve which his budget allows. The maximization and equimarginal principles are behind the scenes, and this chapter tried to show where they were hiding.

The chapter considered some implications of budget-line and indifference curve analysis. It showed that constraints hurt, and that this simple idea leads to the important concept of present value. It looked at the concept of consumers' surplus and how this concept dissolves the paradox of value. Finally, it added the concept of producers' surplus, and showed how fights over the division of surpluses could reduce the total.

QUESTIONS

1. The graph below shows Jane Smith's indifference curves and her budget line. Answer the following questions about points a, b, c, d, e, and f.



a. Which of the points on the graph can Jane afford?
b. Which can she not afford?
c. What points cost the same?
d. Which point gives her the least satisfaction?
e. Which point gives her the most satisfaction?
f. Which point is cheapest?
g. Which point will she choose?

2. What is an isobar? an isotherm? an isotope? an isogram? What do you conclude about most words beginning with "iso"?

3. Ann, Betty, and Carol are three buyers of shoes. Their demand curves are shown below in terms of their willingness to pay.

Number	Marginal Benefit measured as willingness to pay of:		
of Pairs	Ann	Betty	Carol
1	$100	$80	$90
2	20	40	30
3	10	30	10
4	5	10	2
5	2	5	1

a. If the price of shoes is $15, how many pairs will Ann buy? Betty? Carol? How many will be bought in total?
b. What will Ann's consumer surplus be at this price? What will the total consumers' surplus be?

4. You have $10,000. You can invest it for one year at 11%. How much will it be worth in one year? Suppose you will have $10,000 in one year and the interest rate is 11%. What is the present value of this sum?

5. The April 1972 issue of *Consumer Reports* suggested that a home buyer purchase a washing machine, dishwasher, and clothes dryer from a department store which charges $675 plus 15% interest each year for two years rather than having the builder install them. The builder, in this example, could get the appliances for $450, but he would add this amount on to the mortgage. If the mortgage matured in 27 years (the national average in 1972) and if the mortgage rate of interest was 7.75%, the appliances would cost a total of $1075. The total cost of the appliances purchased at the department store on a two-year contract would be $785. Explain why the advice of *Consumer Reports* is bad advice in this case.

6. Draw a set of indifference curves on Figure 5 that indicate a person who would be better off with a cash gift rather than an in-kind gift. b) Can you draw a set of indifference curves representing the preferences of a person who would be as well off with an in-kind gift and a cash gift? What will they look like? c) Can you draw indifference curves which show a recipient who is better off with an in-kind transfer rather than cash?

7. a) Draw a budget line for a person with an income of $200 if good A costs $10 and good B costs $25.

b) Then show what happens to the attainable region if the government gives everyone two units of B.

c) Finally show what the attainable region looks like if the government gives everyone two units of B but forbids anyone to buy more than 7 units of B.

8. Suppose you are given $200 and these two options: A) a sure gain of $50 or B) a 25% chance of winning $200 and a 75% chance of winning nothing. Which would you choose? Now suppose that you were given $400 and these two options: C) a sure loss of $150 or D) a 3/4 chance of losing $200 and a 1/4 chance of losing nothing. Which would you choose? (Problem from Slovic and Lichtenstein, "Preference Reversals," *American Economic Review*, Sept. 1983, p. 601) Most people choose A over B and D over C. Is this consistent with the way the chapter has assumed people to act?

9. John Bauer wants to be a farmer. He can either buy a 500 acre farm which produces 100 bushels per acre, or one of the same size which produces 70 bushels per acre. Explain why the first will not generally be the more profitable for him to buy and cultivate.

10. Your rich uncle gives you and interest free loan of $10,000 for one year. Your rich aunt gives you a gift of $1000 now. Your grandfather promises you $1200, but you will not get it for two years. If the interest rate is 10%, which is the most valuable and which is least? If the interest rate is 5%, will their present value rankings change? Explain.

APPENDIX: INDIFFERENCE CURVES AND MAXIMIZATION

It is not immediately apparent in examining a graph such as Figure 3 that indifference curve analysis contains the maximization principle. To show that it does requires a bit of geometry and algebra, and thus it requires some time and careful thought from the student.

At point c in Figure 3 the indifference curve is tangent to the budget line, which means that they have the same slope. We saw earlier that the slope of the budget line depends only on the relative prices of good *A* and *B*. The formula for the slope is:

(1) Slope of budget line
= -(Price of B)/(Price of *A*).

The slope of the indifference curve shows how easily one good substitutes for the other, and is called the **marginal rate of substitution**. If a sacrifice of one unit of *A* requires a gain of two units of *B* to give the same level of satisfaction, the slope of the indifference curve is -1/2. (See points *a* and *e* on Figure 3. They are on the same indifference curve, so represent the same amount of want satisfaction. To get from *a* to *e*, the person loses one unit of *A* and gains two units of *B*.) This slope means that an extra unit of good *A* is twice as valuable as an extra unit of good *B*, or that the marginal utility of good *A* is twice that of *B*. Thus, the slope of the marginal rate of substitution equals:

$$\text{(2)}\quad \begin{matrix}\text{Slope of an}\\ \text{indifference}\\ \text{curve Marginal}\end{matrix} = -\frac{\text{Marginal Utility of B}}{\text{Utility of A}}$$

Hence at point *c*, the tangency implies that:

$$\text{(3)}\quad \frac{\text{Price of B}}{\text{Price of A}} = \frac{\text{Marginal Utility of B}}{\text{Marginal Utility of A}}$$

which can be rewritten as:

$$\text{(4)}\quad \frac{\text{Marg. Util. of A}}{\text{Price of A}} = \frac{\text{Marg. Util. of B}}{\text{Price of B}}$$

which is the equimarginal principle.

The equimarginal and maximization rules are closely related, but not quite identical. The difference is that the equimarginal principle can be used when the budget constraint is only implicit while use of the maximization principle makes the budget constraint explicit. What this means in practice is that when the maximization rule holds, the equimarginal rule must also hold, but if the equimarginal rule holds, the maximization rule need not hold.

The rules are equivalent with a budget constraint. Then only a limited amount of each can be bought, and the cost of buying more of one is the loss in the other or others that can be bought. If we multiply both sides of equation (4) by the term Pa, we get the equation:

$$\text{(5)}\quad \begin{matrix}\text{Marginal}\\ \text{Utility}\\ \text{of A}\end{matrix} = \frac{\text{(Price of A)(Marg. Utility of B)}}{\text{Price of B}}$$

Since the marginal utility of A represents the extra benefit of A, another way of writing this equation is:

(6) Marginal Benefit of A

$$= \begin{matrix}\text{Number of dollars}\\ \text{needed to buy}\\ \text{one more of A}\end{matrix} \times \begin{matrix}\text{extra benefit of}\\ \text{spending a dollar}\\ \text{on other product}\end{matrix}$$

Table 3 helps illustrate how these equations fit together. Assume that apples costs $.20, bananas costs $.40, and the person has $2.00 to spend. If no apples are bought, then five bananas can be purchased (because five at $.40 each cost $2.00).

Column (2) of Table 3 shows the marginal benefit of additional apples. To get the marginal cost of additional apples, consider what happens when the person buys an apple. By

TABLE 3

(1) Number of Apples or Bananas	(2) Marginal Utility of Apples	(3) Marginal Utility of Bananas	(4) Pa(MUb/Pb) or Marginal Cost of Apples	(5) $\frac{MUa}{Pa}$	(6) $\frac{MUb}{Pb}$
1	20	30	8	100	75
2	19	25	8	95	62.5
3	18	21	9	90	52.5
4	17	18	9	85	45
5	16	>> 16 <<	10.5	80	40
6	15	14	10.5	75	35
7	14	12	12.5	70	30
8	13	10	12.5	65	25
9	12	8	15	60	20
10	11	7	15	55	17.5

the budget constraint, the person must give up one-half a banana, going from five bananas to four and one half. The marginal utility of the fifth banana is 16, so one half of it is eight. This eight is the first number in the fourth column. Alternatively, we could find it from equation 2 as: .20(16/.40)=8. Continuing, obtaining the second apple also requires a sacrifice of one half banana worth 8 utils. When the third apple is obtained, the person must drop from four to three and one half banana, and thus loses one half of 18, or a loss of nine. The rest of the table is constructed in a similar way and shows how much utility is lost when a banana is sacrificed to get another apple. Notice that the table cannot be extended beyond 10 because 11 apples are impossible with the budget constraint.

Using the Marginal Benefit = Marginal Cost rule, we can see that the optimal number of apples is eight. If only six apples are bought, buying the seventh adds 14 units of utility (from column (2)) but costs 12.5 units of utility (from column (4)). Sacrificing 12.5 to get 14 is a good trade and should be made. If an ninth apple is taken, there is a gain of 12 utils from it but the loss of 15 from fewer bananas more than offsets this gain. You can re-compute the solution using columns (5) and (6) and the equimarginal principle. You should discover that the answer is the same.

QUESTIONS

1. Draw indifference curves for the following:

a) right and left shoes (perfect complements).

b. corn from Indiana farms and corn form Illinois farms (perfect substitutes.

c. food and garbage (something godd and something bad).

2. (a) How many indifference curves can exist in a figure such as Figure 3?

(b) Why cannot two indifference curves cross over one another?

CHAPTER 9

THE FIRM AND ITS CONSTRAINTS

After you read this chapter and work the problems at the end, you should be able to:

Define production function, isoquants, marginal product, price discrimination, monopsonist, and the all-or-nothing demand curve.

Define increasing, decreasing, and constant returns to scale.

Distinguish between income and substitution effects.

Distinguish between an individual buyer's demand curve and the industry demand, and between industry demand and the demand curve facing an individual seller.

Compute marginal revenue from the demand curve of the seller when that demand curve is given in the form of a table.

Compute marginal resource cost from the supply curve of the buyer when that supply curve is given in the form of a table.

Explain why marginal resource cost equals price for a buyer who is a price taker.

Explain why marginal revenue equals price for a seller who is a price taker, and why marginal revenue is less than price for a seller who is a price maker.

Explain what the law of diminishing returns is and under what conditions it holds.

Explain why the demand curve, the supply curve for resources, and the production function can be treated as boundaries.

Though economists are interested in many cases of unintended consequences, those unintended consequences which involve businessmen seeking their own gain has been at the heart of economic analysis since Adam Smith. Smith noted that

> "It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interests. We address ourselves, not to their humanity but to their self love, and never talk to them of our necessities but of their advantages."

Since Smith, a great deal of intellectual effort has gone into exploring the question of under what conditions the interests of society will be served by businessmen seeking to make a profit. The next five chapters will present the major conclusions which economists have come to. This chapter begins this presentation by explaining a large number of technical terms which economists use, and also by looking at a few of the simplifying assumptions they generally invoke.

CIRCULAR FLOW

The central interaction of economic life is **exchange**. An exchange is a voluntary agreement between two people. Each gives something to the other and gets in return something which he considers of greater value. When John and Jim exchange baseball cards, John gets cards which he considers more valuable than those he gives to Jim, and Jim gets cards which he considers more valuable than those he gives to John. Unless both sides of the exchange feel that the exchange benefits them, the exchange will not take place. Because both sides benefit, exchange is, in the terms of game theory, a **positive-sum** game.

An alternative to interaction by exchange is interaction involving coercion. With coercion the actions of one side are not voluntary but forced. If Jim takes baseball cards away from John and threatens to beat him up if he complains, we have interaction based on coercion. For now we will ignore interactions based on coercion and consider only those based on exchange. In practical terms, this means we will ignore the government (until Chapter 13), since it is the primary agent which can coerce in any society.

People engage in exchange because the things they get help them attain goals. However, exchange is not just take; in order to get, one must give. People must do things which they do not want to do in order to get things which they desire. The unpleasant part of this process is work and production, and the pleasant part is consumption. Work and production are not pursued for their own sakes, but only because without them we cannot consume. This division of economic life in an exchange can be seen in Figure 1, which shows Frank Knight's Wheel of Wealth, more commonly known as the **circular flow**.

The circular flow diagram divides the economy into two sectors: one concerned with producing goods and services and the other with consuming them. Resources are converted into goods and services by business, and in this transformed state travel back to consumers. Money flows in the opposite direction. These flows involve two markets in which exchange takes place: the resource or factor market in which business buys resources, and the goods and services market in which business sells goods. (Some economists define a "factor of production" as the service of some resource. If resources are land, labor, and capital, the factors of production are the services of land, labor, and capital. We will ignore the distinction between resources and factors of production in the discussion which follows.)

We have already begun to examine how an exchange economy works. Both the model of supply and demand and the discussion of consumer choice with utility analysis are key elements of an understanding of an exchange economy. This chapter looks at the right-hand side of the circular flow diagram, examining the business firm and the constraints or limitations which it must face in its fight for survival. The analysis which this chapter begins is completed in the next chapter, which shows how the rules of maximization

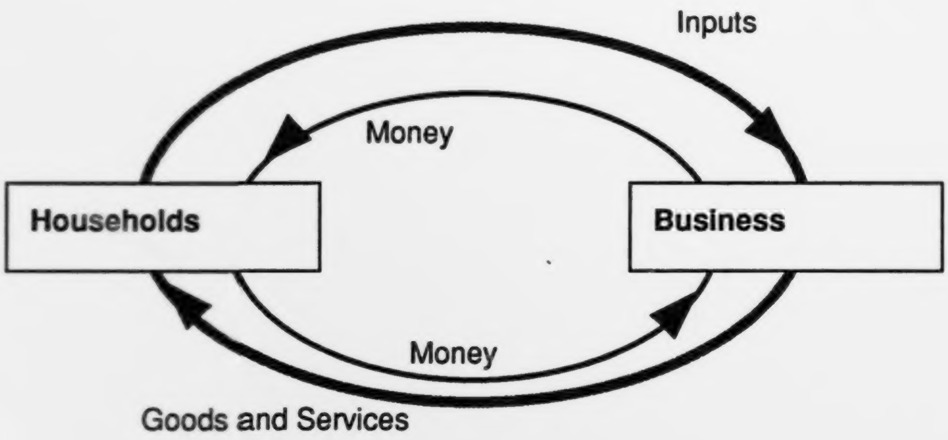


Figure 9.1 Circular Flow

apply to the firm.

DEALING WITH THREE FUNDAMENTAL TASKS

The word "firm" takes on two meanings in economics. Sometimes the word refers to an agent which produces something. Other times it refers to a business organization. The second meaning of the word is more restrictive than the first. A self-employed farmer is not a firm in the second meaning since he is not an organization, nor is a nonprofit organization since it is not a business. Yet both are firms in the first meaning. In this chapter and the next we will usually use the word "firm" in the first sense. Chapter 12 examines the second meaning by asking why people find it advantageous to use organizations to produce.

You should be aware that the word "industry" cannot be used interchangeably with "firm." An industry is defined as all firms producing similar products. It can contain one firm or millions of firms. In practice, there are two complications with this definition. First, there is no clear rule for deciding how similar products must be before they belong in the same industry. For example, what other products belong in the industry which Coca Cola is in? Does Pepsi Cola, or Seven Up, or orange juice, or milk, or beer belong in it? There is no clearly correct answer. Second, the number of firms in an industry depends on the price of the product. Firms enter as price rises and exit as price falls.

The right-hand side of the circular flow diagram, which Figure 2 illustrates, shows three fundamental tasks of all firms in an exchange economy. First, a firm must obtain inputs. Inputs include raw materials, energy, machinery, office space, workers, and anything else needed to produce output. Second, the firm must combine or use inputs to produce output. Output may either be a tangible good such as a pair of shoes or an automobile, or a service such as a haircut or a medical checkup. Third, the firm must sell its output to someone else.

A firm which cannot do these three tasks well enough will not survive. When the auto-

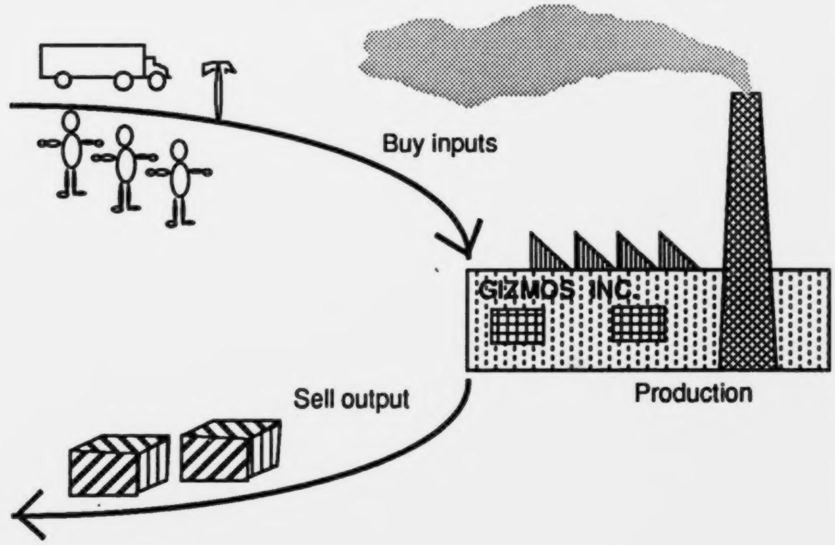


Figure 9.2 A firm's eye view of circular flow

mobile was developed in the early 20th century, firms which made carriages died because they could no longer do task three, selling their output, well enough. Almost all baskets sold in the United States are imported. Baskets are handmade, and no firm in the United States can hire workers at wages low enough to be able to compete with wages which are acceptable in some other countries. Here is a case in which American firms (or firms in any industrialized nation) find it difficult to cope with the first task. The development of the electronics industry is a case in which the second task has changed. New technology allows firms to combine inputs to produce goods which were not possible just a few years ago.

The economic theory of the firm is an analysis of how the firm must perform these three tasks if it is to make a profit. Each task can be described in mathematical or graphical terms. **Supply curves** of resources describe the first task. They indicate how much the firm must pay for the amounts of inputs it wants. The **production function** describes the second task. It tells how much output the firm can produce from a set of inputs. The **demand curve** for output describes the third task. It depends on people's wants or preferences and tells how much the firm can charge for output. Each of these mathematical ways of representing the three fundamental tasks can be seen as a constraint or limitation which the firm faces.

A supply-of-resources curve tells at what prices various amounts of a resource can be bought or hired. Though one can view it in a number of ways, it also can be viewed as a boundary. It tells the firm the *minimum* it can pay for any amount of a resource. Sellers of resources imposed this boundary on the firm which must buy resources in order to produce. Points to the upper left of a supply curve are attainable, while those to the lower right are not.

The production function contains information about how much output can be obtained with various quantities of inputs. The production function is often discussed as a *relationship* between inputs and output, as its name implies. (Mathematically a function is a special sort of relationship.) However, it too can be discussed as a boundary. It shows the *maximum* that can be produced with any combination of resources. Less than this maximum can be produced—one can always get nothing for something.

The demand curve can be viewed from a number of perspectives: as a relationship between price and quantity buyers will buy, as a locus of points of consumer equilibrium, as a measure of marginal benefit to the buyer, or as a boundary. This last view, that the demand curve represents a boundary that buyers impose on the seller, is one that is most useful when developing the theory of the firm. The demand curve limits the amount that sellers can sell at each price. Points to the left of the demand curve are attainable, while those to the right are not.

The remainder of this chapter takes a closer look at these three constraints. The material is somewhat technical, but it forms much of what most economists consider the core of microeconomics. The next five chapters all build on the foundations which this chapter constructs.

THE PRODUCTION FUNCTION

When most people think of fundamental tasks of a firm, they think first of production. Economists describe this task with the production function. The production function is an abstract way of discussing how the firm gets output from its inputs. It describes, in mathematical terms, the technology available to the firm.

TABLE 1
A PRODUCTION FUNCTION

Labor			
5	30	34	37
4	26	30	33
3	21	25	28
2	16	20	23
1	10	13	15
	1	2	3 Capital

A production function can be represented in a table such as Table 1. In Table 1, five units of labor and two of capital can produce 34 units of output. It is, of course, always possible to waste resources and to produce less than 34 units with five units of labor and two of capital, but Table 1 indicates that no more than 34 can produced with the technology available. The production function thus contains the limitations which technology places on the firm.

The production function can also be illustrated in a graph such as Figure 3. This graph looks exactly like a graph of indifference curves because the mathematical forms of the production function and the utility function

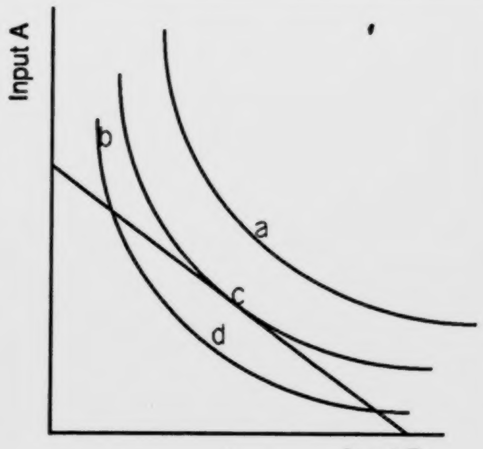


Figure 9.3 Isoquants show a production function graphically.

are identical. In one case, inputs of goods and services combine to produce utility, and in the other, inputs of resources combine to produce goods or services. A curved line in Figure 3 shows all the combinations of inputs that can produce a particular quantity of output. These lines are called **isoquants**. As one moves to the right, one reaches higher levels of production. If one can visualize this as a three dimensional graph, one can see that the production surface rises increasingly high above the surface of the page; the isoquants indicate a hill. The firm must operate on or below this surface.

There is one rule that seems to hold for all production functions, and because it always seems to hold, it is called a law. The **law of diminishing returns** says that adding more of one input while holding another input constant results eventually in smaller and smaller increases in added output. To see the law in Table 1, one must follow a column or row. If capital is held constant at two, the marginal output of labor (which economists usually call **marginal product** of labor) is shown in Table 2 below. The first unit of labor increases production by 13, and as more labor is added, the increases in production gradually fall.

The law of diminishing returns does not take effect immediately in all production functions. It is possible for the first unit of labor to add only 4 units of output, the second to add 6, and the third to add 7. If a production function had this pattern, it would have increasing returns between the first and third worker. What the law of diminishing returns says is that as one continues to add workers, eventually one will reach a point where increasing returns stop and decreasing returns set in.

TABLE 2
THE MARGINAL PRODUCT OF LABOR

LABOR	MARGINAL OUTPUT
FIRST	13
SECOND	7
THIRD	5
FOURTH	5
FIFTH	4

The law of diminishing returns illustrated in Table 2 is not caused because the first worker has more abilty than the second worker, and the second is more able than the third. By assumption, all workers are the same. It is not ability which changes, but rather the environment into which workers (or any other variable input) are placed. As more and more workers are added to a firm with a fixed amount of equipment, the equipment must be stretched over more and more workers. Eventually the environment becomes less and less favorable to the additional worker. People's productivity depends not only on their skills and abilities, but also on the work environment they are in.

The law of diminishing returns was a central piece of economic theory in the 19th century and accounted for economists' gloomy expectations of the future. They saw the amount of land as fixed, and the number of people who could work the land as variable. If the number of people expanded, eventually adding one more person would result in very little additional food production. And if population had a tendency to expand rapidly, as economists thought it did, one would predict that (in equilibrium) there would always be some people almost starving. Though history has shown the gloomy expectations wrong, the idea did have an influence on the work of Charles Darwin, so it was not without some redeeming merit.

If one increases all inputs in equal proportions, one travels out from the origin on a ray. There is no law to predict what will happen to output in this case. If a 10% increase in all inputs yields more than a 10% increase in output, the production function has **increas-**

ing returns to scale. If it yields less than a 10% increase in output, the production function has **decreasing returns to scale**. And if it yields exactly a 10% increase in output, it has **constant returns to scale**.

Returns to scale are important in determining how many firms will populate an industry. When increasing returns to scale exist, one large firm will produce more cheaply than two small firms. Small firms will thus have a tendency to merge to increase profits, and those which do not merge will eventually fail. On the other hand, if an industry has decreasing returns to scale, a merger of two small firms to create a large firm will cut output, raise average costs, and lower profits. In such industries many small firms should exist rather than a few large firms.

Most products require many more than two inputs, but showing a production function with more than two inputs with graphs or tables is difficult. Products require various types of labor and capital, energy of various sorts, and raw materials. One of the key inputs, especially in larger firms, is managerial ability. Inputs do not combine by themselves to produce output. Someone must have knowledge of how to combine inputs and to coordinate the production process.

If business decision makers lack information or are incompetent, the firm will not make the best use of available resources. Or if morale is bad in a firm, people may work poorly and produce less than they could. In either case the firm will produce below the maximum that the production function allows. Economist Harvey Liebenstein has called losses of these sorts "X-inefficiency." Though economists assume that the firm will be on the production function, a major challenge of management is to make decisions so that the firm will be on or close to the production function.

In explaining the theory of the firm, economists conventionally assume that the production function is fixed and that the firm operates on the surface of the production function. The firm need not consider the production function as fixed, but may view it as a variable that it can alter through research and development. Creativity in the form of new technology or new management techniques may loosen the boundary that the production function represents and may make possible greater profit, at least temporarily.

THE DEMAND CURVE FOR OUTPUT

Once a firm has produced a product, it must sell it. The demand curve for output describes the limitations the firm faces in doing this task. The demand curve for output is a constraint on the firm because it gives the maximum price that a firm can charge if it wants to sell some amount of its product. If the firm wants to sell 24 in Figure 4, it can do so by charging $5.00 or any price that is lower. It cannot charge $10.00 and still sell 24 because buyers will not let it.

The demand curve facing a firm depends both on the preferences of consumers and on how well other firms meet those preferences. Chapter 8 showed how one could derive a demand curve for an individual from a set of indifference curves showing the individual's

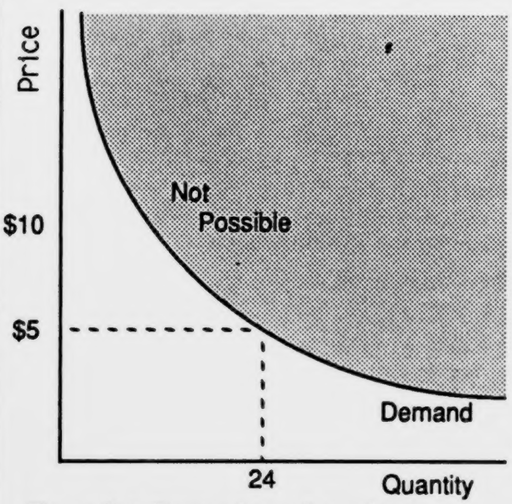


Figure 9.4 Demand as a boundary.

preferences and a series of budget lines showing changes in price. To get a demand curve for the entire industry, one must add up all the demand curves of individuals. To get the demand curve for eggs, for example, one must add up the number of eggs that Smith and Jones and Nelson and all other consumers in the market want at each possible price.

When there is only one firm selling in a market, that firm is a **monopolist**. (The Greek root *mono-* means "one.") The demand curve for the monopolist is the demand curve for the industry. A monopolist is a **price searcher** or a **price maker**. It will search along the demand curve for the price-quantity pair which is most profitable. When there is more than one seller, the demand curve which a seller sees is not the same as the demand curve for the industry. The industry demand is split up among sellers. When there are only a few sellers, the sellers will still be price searchers or price makers. These sellers, or **oligopolists**, (the Greek root *oli-* means "few") are price makers because each recognizes that if it wants to sell more, it must lower its price.

However, the demand curve of each oligopolist will be more elastic than the demand curve for the industry as a whole. Suppose, for example, that there are two firms in an industry, that each produces 50 units of output, and the elasticity of the industry demand curve is one. If one firm increases its output by 10% to 55, the industry output increases to 105, which is a 5% increase. Since the price elasticity of demand is one, price must decline by 5%. But for the original firm, a 10% increase in production and a 5% decline in price indicate a price elasticity of two, not one.

As firms get more and more numerous in an industry, the demand curve each sees gets more and more elastic. When there are a great many sellers in the market, a change of output by any one of them has an insignificant effect on price. To each firm the demand curve will look perfectly flat—the firm will seem able to sell whatever amount it wants at a fixed price. In this case each firm is a **price taker** and sells in a perfectly competitive market. An example of this type of market is the market for wheat. There are a great many wheat farmers in many countries, and none has any noticeable control over the price at which it can sell in the world wheat market.

However, even when there are a great many sellers, each firm may have a downward-sloping demand curve. If buyers must expend time and effort to discover prices or the characteristics of the product, they will pick a seller and stay with it as long as they find the exchange satisfactory. These downward-sloping demand curves of small sellers are a result of the ambiguous definition of industry. The products most firms produce differ in some way, such as in quality, service, or location, from the products of other firms in the industry.

From the viewpoint of the firm, it is not the demand curve, but the child of the demand curve, the **marginal revenue** curve, which is of vital importance. You have already seen in Chapter 5 that marginal revenue is the extra revenue a seller gets when it produces and sells another unit. Marginal revenue depends on the demand curve and is derived from it. For the price taker, the marginal revenue curve is the demand curve. For the farmer who can sell corn at $4.00 a bushel, the extra revenue from selling another bushel is $4.00. The demand curve for this farmer is flat at $4.00, and so is his marginal revenue curve.

However, for a price maker marginal revenue will be less than price. Table 3 helps show why. If the firm charges $3.00, it can sell one unit and total revenue will be $3.00. If it sells one more unit, it will be forced to cut price to $2.00 and total revenue will rise to $4.00. Selling the extra unit adds only $1.00 to sales. Though the second unit sold for $2.00, the firm had to cut the price it was previously receiving for the first unit by $1.00, so the net increase in revenue was only $1.00. By similar logic, selling the third unit reduces total revenue by $1.00, so marginal revenue is -$1.00.

The above analysis assumes that the firm can charge only one price. If it can charge more than one price, charging higher prices

TABLE 3
DEMAND AND MARGINAL REVENUE

PRICE	QUANTITY	MARGINAL REVENUE
$3.00	1	
		$1.00
$2.00	2	
		-$1.00
$1.00	3	

to those willing and able to pay them and lower prices to others, it can move the marginal revenue curve closer to the demand curve. This move will increase profits (or reduce losses). This pattern of pricing, called *price discrimination*, is discussed in Chapter 15.

We have discussed the demand curve as a boundary, but a special case of price discrimination may occasionally allow a seller to get to the right of the demand curve to another curve called the "all-or-nothing demand curve." Suppose the seller in Table 3 tells the buyer that he has a choice, either nothing or three units for $1.90 each. The buyer would be better off buying three rather than none because the total value of the three is $6.00. Recall that the demand curve tells the marginal benefit to the consumers. The first unit is worth $3.00, the second is worth an added $2.00, and the third one is worth an added $1.00, for a total of $6.00. At $1.90 each, the three cost $5.70, so the buyer will buy them. He would prefer only two at $1.90 each, but that option is not open to him.

It is very difficult to find examples of firms trying to reach the "all-or-nothing" demand curve. This behavior is most likely to occur only in markets with one seller and few buyers, markets in which prices are set in a bargaining process.

Economists generally assume that the demand curve is fixed, but many businesses do not regard it that way. It can vary seasonally, with the general level of business activity, or with a trend. The demand for turkeys has a pronounced seasonal movement. The demand for automobiles changes when there is a recession. The demand for baby food follows the trends in birth rate.

Business also may be able to move its demand curve through advertising. Advertising may simply give people information, it may change their goals, or it may change their perception of the product. For the firm it does not matter which happens. The result is the same—good advertising moves the demand curve to the right.

The demand curve can move for other reasons. If the firm lowers its price, and then raises it back to its previous level, it may find that sales at the old price have changed. The lower price may attract new customers who have not tried the product before, and who find they like the product enough to stick with it when the old price is restored. Alternatively, some customers may expect prices to be cut again sometime in the future, and may decide to postpone purchases until it happens again. The opposite can happen if the firm temporarily raises price. It may encourage some customers to try substitutes, which they may find suit them better than the original product. Or it may encourage customers to buy more when the price comes down to prepare for any future increase.

The firm may also be able to change its demand curve by changing the characteristics of its product.

Finally many firms sell several products that may be interrelated, and any pricing decision on one product will have effects not only on that product but on others. For example, the prices General Motors charges for Chevrolets will affect the demand curve for Pontiacs.

SUPPLY OF RESOURCES

The third task of the firm is to obtain resources needed to produce a product. For each resource, a supply curve shows limitations which the firm faces. A supply curve for a resource is based on the preferences of sellers and on the actions of other firms which use the resource.

We have already seen how a market demand curve can be derived from utility curves and a budget line, and it may seem surprising that a supply curve can also be derived from the same procedure. To see that it can, consider Figure 5, which shows indifference curves for income and leisure. Income is desirable because one can obtain other desirable things with it, and leisure is desirable because it lets one enjoy income.

In the world we live in, greater amounts of income must be purchased with more work—which means less leisure. The tradeoff between leisure and income will depend on the budget line, which depends on the wage rate. If the wage rate is $10, some of the options open to an individual include 24 hours of leisure and no income, or 20 hours of leisure and $40 of income, or 10 hours of leisure and $140 of income.

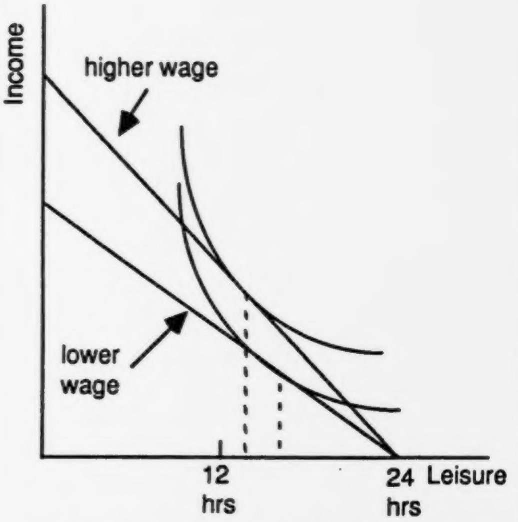


Figure 9.5 Finding a labor supply curve.

To get a supply curve for labor, one must see what happens if the wage rate changes. Two wage rates are shown in Figure 5. At the higher wage rate, the individual wants less leisure (which means he will work more as wages rise), but one could as easily draw indifference curves which show that as wages rise, the amount of leisure rises instead of falls. Higher wages have two effects on the leisure-work decision, and these two effects pull in opposite directions. A higher wage rate increases the benefits of working, causing people to substitute work for leisure. This is called the **substitution effect** and is caused by changes in the slope of the budget line. Higher wages also increase income, and with a higher income, people want more leisure. This is called the **income effect** of a price change, and is caused by changes in distance of the budget line from the origin.

Usually the substitution and income effects re-enforce each other, but in the case of labor they pull in opposite directions and they about cancel out. Most economists believe that the market supply curve for labor (found by adding up all the supply curves of individuals) is close to a horizontal line.

Supply curves for other resources can be obtained in similar ways. The supply curve for capital, for example, depends on decisions of people to consume now or to consume in the future. People who prefer to consume in the future will save, and make funds available to finance capital. Their "time preference" determines the shape of indifference curves. The slope of the budget line depends on the interest rates. The budget line tells how much one can get in the future if one sacrifices consumption now. Question 6 at the end of the chapter develops this topic further.

Though the market supply curve for labor may be almost vertical, few firms see this supply curve. If many other firms also are buying labor, what one firm does may have little effect on the overall market. If the firm is so small in the market that it can see no effects at all on the wage from its hiring decisions, it is a price taker. If it has some effect on wages, so that when it wants to hire more, it finds that wages rise, the firm is a price maker. The extreme case of a price maker is a **monopsonist**, the case of only one *buyer* in the market. The supply curve for a resource which a monopsonist sees is the same as the market supply curve.

The supply curve for a resource is a constraint or boundary on the firm because it shows the minimum that the firm can pay for a level of the resource. If the firm is a price taker in the resource market, it will face a horizontal supply curve such as that in Figure 6. This curve indicates that any number can be bought at P1 with no effect on price. There is no way the firm can attain point *a* even though it might prefer to pay less than P1 because no one will sell at less than P1. Sellers will not sell because we assumed that there were a great many other buyers of the resource who will pay P1. Point *c* is possible, but a waste of money since the same amount of the resource could be bought for less.

If the firm is one of a few buyers or the only buyer of a resource, then it may face a supply curve which slopes upward. In this case the buyer is a price maker or price searcher. It can obtain quantity Q1 if it pays P1, but if it wants quantity Q2, it must pay more than P1.

You should have noticed that there are similarities between a supply curve for a resource and a demand curve for output. Both are boundaries, and the curve the firm faces may differ from the market curve. The similarity goes further, because there is a counterpart for marginal revenue. It is **marginal resource cost**. Marginal resource cost measures the extra cost to the firm of hiring one more unit of the resource.

When a firm is a price taker, marginal resource cost is the same as the price of the resource. If the firm can hire as many workers as it wants at $10 per hour, then hiring one more hour of labor adds $10 to costs. Marginal resource cost and the supply of labor are both horizontal lines in this case.

When a firm is a price maker, which means it faces an upward-sloping supply curve, the extra cost of hiring another unit of the resource is different from the price of the extra unit. Table 4 illustrates the reason for this difference. In Table 4 the price of one unit is $1.00, the price of two units is $2.00 for each, and the price of three units is $3.00 for each. Thus if the firm wants to buy two units, it cannot pay $1.00 and get two. It must be willing to pay $2.00. However, the added cost of the second unit is not $2.00, but $3.00.

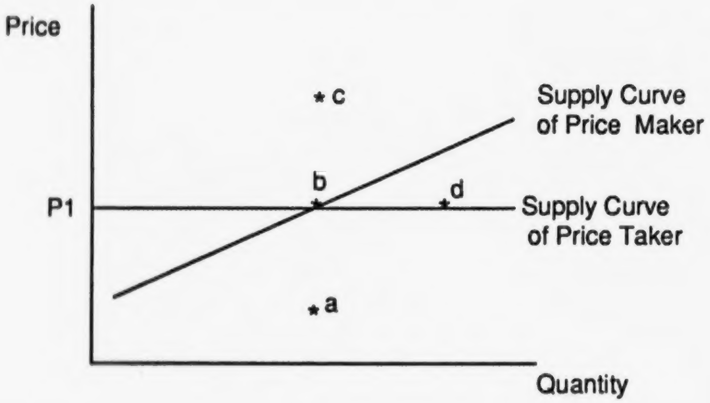


Figure 9.6 Supply curves for resources.

TABLE 4
COMPUTING MARGINAL RESOURCE COST

QUANTITY	PRICE	MARGINAL RESOURCE COST
1	$1.00	$1.00
2	2.00	2.00 + 1.00
3	3.00	3.00 + 2.00

This can be shown by comparing the total costs of two units and one unit, or $4.00 less $1.00. The added cost of the second unit is not only the two dollars that must be paid for it, but an added dollar for the first one. By the same logic the added cost of a third unit is $5.00.

The marginal resource cost curve lies above an upward-sloping supply curve because of the assumption that the firm can pay only one price. This is often a realistic assumption. If the firm hires two clerks who do exactly the same work, and pays one $4.00 per hour and the other $6.00 per hour, the lower-paid one will be unhappy and may refuse to work for the lower pay. On the other hand, many firms do not publicly disclose what they pay various people, and discourage employees from discussing salaries. To the extent that people are unaware of what others are earning, the firm may be able to pay different prices for the same resource, or in economists' jargon, to price discriminate. Price discrimination will pull down the MRC curve in Figure 7 closer to, or perhaps even onto, the supply curve.

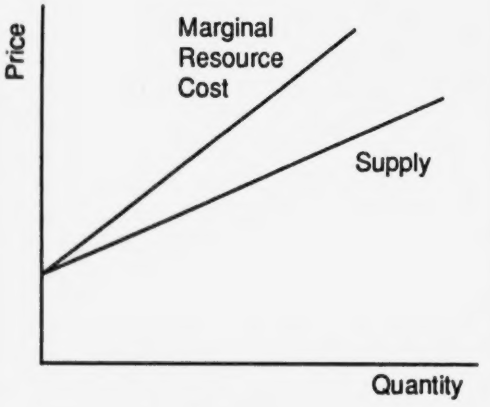


Figure 9.7 Marginal resource cost when supply slopes upward.

SUMMARY

This chapter introduced the economic theory of the firm which the next chapter will complete by showing how the rules of maximization apply to the firm. The major theme of this chapter has been that in order to make a profit, a business must deal with three constraints or boundaries: the demand curve for output, the production function, and supply curves for inputs.

Table 5 shows how these concepts are related to the functions of the business and to the marginal concepts which will be central in the next chapter.

The importance of these three constraints can be seen in a product which no firm can produce at a profit. Any one of three changes, if large enough, can change the situation and make a profit possible. The cost of the inputs may drop in price enough so that the product is profitable. Or technology may improve enough to make the product profitable. Or people may increase the amount they are willing to pay for the product enough so that it is profitable. These three changes are changes in the supply-of-resources curves, the production function, and the demand curve, respectively.

TABLE 5
THE CONSTRAINTS FACING THE FIRM

Task:	Economic Concept	Limitation Imposed by	Relationship between:	Marginal Concept
buying inputs	supply of resources	resource owners	price (money) and amount of resource	marginal resource cost
producing output	production function	technology	inputs and output	marginal product
selling output	demand curve	customers	output and price (money)	marginal revenue

QUESTIONS

1. Fill in the blanks:

Amt of Capital	Total Cost	Output	Total Revenue
20	$50,000	89	$55,000
22	$50,100	90	$55,300

a) Increasing capital by 1 increases costs by __________
b) Increasing capital by 1 increases output by __________
c) Increasing output by 1 increases revenue by __________
d) Increasing output by 1 increases cost by __________
e) Increasing capital by 1 increases revenue by __________

Answer the same questions using the table below and the technical terms of economics:

Amt of Capital	Total Cost	Output	Total Revenue
109	$120,000	500	$130,000
111	123,000	505	134,000

f) Marginal resource cost of capital = __________
g) Marginal product of capital = __________
h) Marginal revenue = __________
i) Marginal cost = __________
j) Marginal revenue product = __________

Fill in the blanks:

Amt of Capital	Total Cost	Output	Total Revenue
100	$1000	100	$1000
____	______	101	____

a) Marginal resource cost of capital = $10
b) Marginal product of capital = 2
c) Marginal revenue = $6
d) Marginal cost = 5
e) Marginal revenue product = ______________

2. Students often confuse production functions and production-possibilities frontiers. They are different relationships, but they are related. The following exercise illustrates their relationship.

Suppose an economy has three units of capital and four units of labor. It must use these resources to produce the only two goods it uses, food and shelter. The technology possessed in this economy is shown in the following production functions:

		FOOD			
Labor	4	7	27	31	34
	3	5	21	25	28
	2	3	16	20	23
	1	1	10	13	15
	0	0	0	0	0
		0	1	2	3

	SHELTER			
4	0	4	5	6
3	0	3	4	5
2	0	2	3	4
1	0	1	2	3
0	0	0	0	0
	0	1	2	3
				Capital

Notice that if resources are fixed in amount and if they are used to produce food, they cannot be used to produce shelter. If, for example, two units of capital and two units of labor are used to produce food, then only one unit of capital and two units of labor will be available to produce shelter.

Complete the production possibilities table for this economy:

Shelter	Maximum Food That Can Be Produced
0	34
1	
2	
3	
4	
5	
6	0

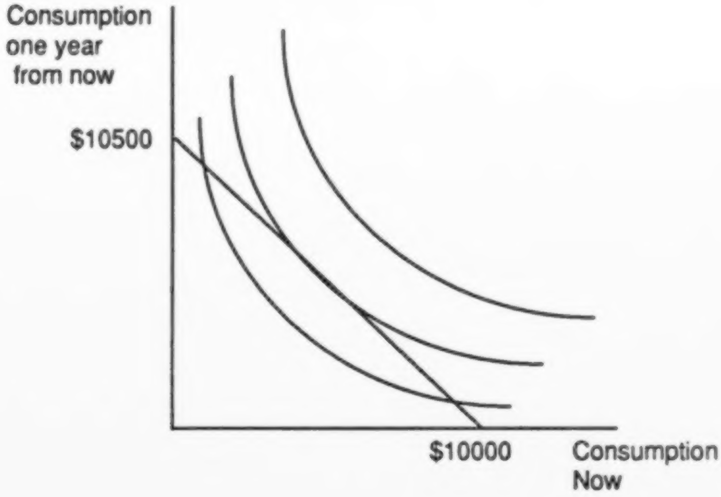
The only two changes which alter a production-possibilities frontier are changes in technology or resources. Why?

3. Many firms sell their output to other firms, which then process it further. For example, the steel industry sells steel to automobile companies. What do transactions of this sort look like on a circular-flow diagram?

4. If any input is held fixed, the marginal cost curve will slope upwards because of the law of diminishing returns. Explain why.

5. If a firm holds excess capacity to discourage entry of new firms, is it on the surface of its production function?

6. The graph below shows a person's indifference curves for consuming now or in the future (one year from now).



a) Suppose the person will get paid $10000 now and nothing else. How can the person obtain the $10500 one year from now?
b) What is the interest rate which the budget line shows?
c) Suppose the person will receive $10500 one year from now and has nothing now. How can consumption take place in the present?
d) What will happen in the diagram if the interest rate increases to 10%?
e) If the interest rate increases, postponing consumption is more rewarded and savings should increase. However, many economists argue that savings may not increase, and in fact a higher interest rate may cause people to consume more in the present, or save less. What force do they rely on to argue that savings can decrease with a higher interest rate? (Hint: Look back and see the forces involved in the effects of wage rates on labor. If people consume more now with a higher interest rate, must they consume less later?)

7. Show a budget line facing a worker who is paid $5.00 per hour. (Hint: look at Figure 9.5.)

a) Draw a set of indifference curves so that he wants to work exactly eight hours at this wage.

b) Draw in a new budget line for a wage of $10.00.

c) Now draw in a budget line which shows him paid $5.00 per hour for the first eight hours, and $10.00 per hour after that. Is this budget line or that in part b more likely to get him to work longer hours? Explain in terms of the income and substitution effects.

CHAPTER 10

MAXIMIZING PROFIT

After you read this chapter and work the problems at the end, you should be able to:

Explain why economic profit is often impossible to measure.

Explain the principle-agent problem.

Defined fixed costs, marginal revenue product.

Compute marginal revenue product when given marginal revenue and marginal product.

When given data on output, marginal revenue, and marginal cost, determine which level of output maximizes profit.

When given data on levels of input, marginal revenue product, and marginal resource cost, determine the level of input which maximizes profit.

Show how the two ways to maximize profits (by looking at optimal output and optimal input) uses the information from the three constraints a firm faces (the production function, the demand curve, and the supply curve of resources).

The business firm is the productive unit in an exchange economy. In the last chapter we saw that in order to survive, a firm had to deal with three constraints: the demand for its product, the production function, and the supply of its inputs. When the firm successfully deals with these constraints, it makes a profit.

This chapter explores the conditions under which a firm maximizes profits. The assumption of profit maximization lets us see how the rules of maximization developed in Chapter 7 apply to firms. The chapter also considers the question of whether the assumption that firms do maximize profits is a good assumption.

PROFIT

Economists conventionally assume that firms attempt to maximize profit. This assumption has been a source of controversy, in part simply because the word "profit" is a bit nebulous. One might think that the word "profit" is well defined because each year thousands of firms announce to the public exactly what their profits are. Firms compute these profits by subtracting from their total receipts most cash outlays plus an allowance for depreciation of capital. (This is the accounting definition of profit greatly simplified. There are in fact many technicalities that complicate this computation, but they need not concern us here.) Profit emerges as a residual—something left over after costs have been paid.

The problem that economists have with this definition is that it ignores *implicit* returns. The ma-and-pa grocery or restaurant may seem to have a profit when costs are subtracted from revenues, but this may be only because the owners do not pay themselves a wage. When their time is valued at even low levels, the "profit" may disappear. Or the family farm may have a profit, but only because it neglects to take into account a return on the land. If this return is computed on the basis of what rent could be obtained on the land, and if an allowance is made for the value of the farmer's labor, there may be little or no profit. Much of corporate profit can reflect an implicit return on investment. If the corporation has a large investment in capital and this entire investment was financed by borrowing, then the return on capital is captured in the interest payments the firm makes. But if the investment is not financed by borrowing, then the return on it will be reported as profit in the income statement.

The economic definition of profit, then, is the difference between revenue and the *opportunity cost* of all resources used to produce the items sold. This definition includes implicit returns as costs. Since profit is a surplus in this definition, it should not exist in industries where entry is easy. Whenever a surplus exists, new firms should flow into an industry, bidding up the price of resources and bidding down the price of output until profit in the economic definition is eliminated. Profit should not exist in long-run equilibrium.

Profit does exist in the real world, and there are several explanations economists have for it. Some economists simply consider it as an indication that the economic system is in perpetual disequilibrium. Joseph Schumpeter, for example, saw profit as a return to a successful entrepreneur. The entrepreneur who finds an opportunity where no one before him saw one, and takes advantage of this opportunity, will make a profit. However, this profit will be temporary, because as time goes on others will follow him and erode his profit. Others see profit as an indication that forces of competition may not be strong, and that in some industries barriers to entry exist. Still other economists have argued that profit is a special sort of implicit return, a return for bearing risk. Those who are willing to take more risk will, on the average, earn higher returns.

Once one includes implicit costs as part of costs, an accurate measurement of profit is impossible, which is a major reason that accountants do not try to measure the economic concept of profit. This obviously causes problems if one wants to test whether or not firms try to maximize profits. An alternative way to approach the measurement problem begins with the Schumpeterian notion of profit, which is that it comes from the discovery of opportunity. A discovery of an opportunity should increase the discover's wealth. Thus a way to measure profit is to try to measure unexpected changes in wealth—wealth increasing faster than a normal rate of return. If one sees people whose wealth is increasing more rapidly than it would if their assets were entirely in the form of high-grade, short-term bonds, one could conclude that they had found a profitable opportunity. There are measurement problems here as well, however. Much of people's assets are in the form of investment in themselves—human capital—and is not easily measured.

Do firms try to maximize profits? If profit is defined as accountants measure it, it seems unlikely that firms do. However, if costs are defined broadly enough, then, since profit is good for the firm, it must be maximized if the decision maker is rational. For example, suppose a humanitarian opened a business with a goal of helping the poor. He might keep prices below a level at which accounting profits are maximized. Yet one can argue that if the humanitarian raises prices, his (opportunity) costs increase because he must partially forgo the goal of helping the poor. Thus with a loose enough definition of cost, one can argue that a very wide assortment of observed results are consistent with the assumption of profit maximization. (However, with a very loose definition of costs, the hypothesis may cease being a scientific hypothesis since it may not be open to refutation.)

Even if there were no definitional problems, there still would be a fundamental problem with the assumption that firms maximize profits. Most firms do not have a single decision maker. Rather they are made up of an assortment of individuals, each using the firm to attain his own goals. Managers pursue goals, stockholders pursue goals, blue and white collar workers pursue goals, but the firm does not.

The assumption that *firms* attempt to maximize profits is inconsistent with the underlying methodology of microeconomics, which assumes that all decision makers are individuals. Mainstream economists criticize Marxist economists for dealing in terms of class actions without grounding these actions on individuals pursuing goals. Given that most production takes place in firms which are large organizations, the assumption that firms attempt to maximize profits is open to the same criticism that mainstream economists make against the Marxists. The assumption that firms maximize profits also needs to be grounded on individuals pursuing goals.

One attempt to ground the assumption of profit maximization on individual behavior appeals to the **principle-agent problem**. An agent acts on behalf of another person, called a principle. How can the principle be sure that the agent acts in the principle's interest? The principle must restrict the agent in some way so that it is in the agent's interest to act in the interest of the principle. The most common way to do this is to tie rewards with performance.

Payment on a commission basis is an important way to solve the principle-agent problem. When one hires a lawyer to sue on a civil matter, or an auctioneer to sell one's belongings, or a real-estate agent, or when a movie star hires an agent to seek employment, payment is done with commissions. When the agent does a good job, he or she is paid a lot. When the agent does a poor job, he or she is paid little or nothing.

Sometimes the principle can protect himself with a contract which specifies how the task is to be performed and what the price will be. The ability to use the courts to enforce the contract protects the principle. Sometimes the principle can protect himself by a threat to take the task elsewhere. This threat is effective is repeat business is important to the agent.

The principle-agent problem is one of control: how does the principle control those working on his behalf. Once one starts looking for cases of this problem, one finds them everywhere. How do shareholders control management, or citizens their government, or management their employees? The principle-agent problem is everywhere because we live in a world of interdependence and specialization.

The principle-agent problem occurs in several places in the firm, including in the relationship between legal owners and hired managers. Stockholders own shares because they believe that this ownership will increase their wealth. Managers work for the firm primarily because it provides them with income needed to buy goods and services, and also because their positions provide them with prestige and authority. The threat of a takeover or hostile merger provides stockholders with their most effective constraint on managerial action. If managers perform poorly at making profits, shareholders will sell their shares because they will not achieve their goal of increasing wealth. The lower value of the company is an inducement for other management groups to buy out the company and replace management. If the new managers can improve performance, they can capture at least some of the increased value of the firm.

In this chapter we will assume that the firm does maximize profit. The assumption has become standard ever since economics developed the theory of the firm in a mathematical framework. The assumption gives the theory an elegance and simplicity that would otherwise not be possible. Before we do this, however, let us take a look at a different assumption, that organizations simply try to survive. The assumption that members of the firm want the firm to survive may be a more realistic assumption than the assumption that firms attempt to maximize profits. The continued existence of the firm provides benefits to owners, managers, and workers. Members of the firms can and do try to eliminate other members whose actions threaten the survival or health of the firm. Since survival of the firm requires profit, the assumption that members of the firm want the firm to survive is equivalent to an assumption that the firm attempts to make a profit. Further, as we will see later in this chapter, the assumption of survival can, under some conditions, lead to profit maximization.

The assumption of survival reveals similarities between nonprofit organizations and for-profit organizations. It is not hard to see some of the similarities. A not-for-profit organization faces the same three constraints that a for-profit organization does: supply curves for resources, a production function, and some sort of demand curve for output. For example, many art museums exist as nonprofit, private organizations. They must buy resources (hire guards and curators, pay utility bills, etc.), they produce a product (a service of art exhibition), and they must obtain money in order to continue in existence (donations and entrance fees).

The major difference between the for-profit organization and the nonprofit organization is that the former announces that its goal is to make a profit, and then is forced by the constraints to produce a product that is valuable to someone, while the latter announces that its goal is to produce a product that is valuable to someone, and then is forced by the constraints to cover its costs. But if an organization covers its costs, its revenues must exceed costs, which means it is making a profit.

One could imagine an art museum organized as a for-profit corporation. One would expect it to charge an admission fee that would maximize profits, and one expects this to be higher than the fees a not-for-profit museum would charge. A not-for-profit museum wants to encourage art enjoyment, and higher fees interfere with this goal. But then why do museums charge any fees at all? A zero fee would maximize the number of visitors and art enjoyment (ignoring problems of congestion). At one time few museums charged fees. Now most museums could not survive without fees. Since both nonprofit and for-profit organizations face the same three constraints, the actions of one can be hard to distinguish from those of the other if the constraints allow only a small area of profit. As a result, the economic theory of the firm describes, with a few minor modifications, the options open to not-for-profit organizations. (It is interesting that a for-profit museum would be at a competitive disadvantage to a not-for-profit museum since the not-for-profit status encourages donations, an important source of revenue.)

THE PROFIT-MAXIMIZING OUTPUT

To determine which level of output to produce becomes a mathematical problem when all information is known about the three constraints which the firm faces: production function, supply curves for resources, and demand curves for output. When they are not known, the problem of maximizing profit is often solved by trial and error, though an understanding of how a solution is obtained in the mathematical model with perfect information may help guide the trial-and-error process.

The maximization principle tells us that the profit maximizing level of output will be that level at which the marginal benefit to the firm of producing one more unit will just equal the marginal cost of that unit. The marginal benefit to the firm is the extra revenue from producing and selling another unit of output. This extra revenue is marginal revenue and depends only on the demand curve. The marginal cost is more complex, because it depends both on the production function and the supply of resources. To discuss it, one must first compute a total cost curve for output.

TABLE 1
A PRODUCTION FUNCTION

LABOR			
5	31	35	38
4	27	31	34
3	21	26	29
2	16	20	23
1	10	13	15
	1	2	3
			CAPITAL

Total cost depends on the prices of inputs and the output that those inputs can produce. Suppose labor costs $12 and capital $5. Then to produce 27 units of output in Table 1 with four units of labor and one unit of capital would cost $53. However, there is a cheaper way to produce 27 because 29 can be produced with three units of labor and three units of capital for $51.

There are many ways to produce most products. In the United States farming is done with a lot of equipment and only a few people. In China it is done with many people and few machines. A hole can be dug for a skyscraper with lots of labor and little equipment, or little labor and a lot of equipment. Electricity can be produced from a dam which uses no fuel but a very large amount of capital, or from turbines which use a lot of fuel but little capital. However, not all of these various ways of doing things are equally cheap. What seems advantageous in one situation may not be in another where the cost of resources is different.

To find the total cost curve, we need to find the cheapest way to produce 27 units of output, 28 units of output, and all other levels of output. We begin by searching the production function for all the ways to produce 27 units of output, and then compute the cost of each of those ways. Only the cheapest of those ways will be on the total cost curve. Then we must do the same for 28 units of output, and for all other amounts of output. When we have found the cheapest ways to produce each level of output, we will have found the total cost curve.

The total cost curve will slope upward, as in Figure 1. It divides a region which is attainable from a region which is not attainable. Point *a* is attainable, while point *b* is not. Point *a* represents an expensive way to produce output. The fact that it is above the total cost curve means that a cheaper way to produce the same amount of output exists. Point *b* represents a method of production which does not exist. It is below the cheapest existing way to produce output.

Every point on the total cost curve in Figure 1 will satisfy the equimarginal principle of:

$$MP_A/MRC_A = MP_B/MRC_B = MP_N/MRC_N$$

In words, the equimarginal principle says that ratios of marginal products of each resource to marginal resource cost of each resource must all be equal. Recall that the marginal product is the amount of extra output that an extra unit of the input would produce (this is a rough definition—the correct mathematical description is in terms of derivatives, a calculus concept)—and that the marginal resource cost of each resource is how much the firm must pay to get one more unit of the resource. If the ratios are equal, there is no change that could get more output at a lower cost. If it is not true, then costs can be cut with output held constant, or output can be increased with costs held constant.

A numerical example can show that if the equimarginal principle does not hold, a firm can increase output without changing costs. Suppose the marginal product of resource A is 10 and the firm adds $5.00 to its costs to obtain a unit of this resource. The marginal product of resource B is 9 and the firm adds $1.00 to its costs to obtain another unit. Putting these numbers into the equation above, we see that 10/5 is not equal to 9/1. If $5.00 were shifted away from resource A, the firm would lose 10 units of output, but would gain an additional 45 if the money were used to buy more of resource B. This would give it a net gain of 35. As money is shifted, the marginal products will change until the ratios are brought into equality.

Notice that the equimarginal principle requires information from both the production function (the marginal product) and the supply curves of resources (the marginal resource costs). If the firm is a price taker in all resource markets, the marginal resource costs will be the same as the price it pays. (You may notice that one can present the discussion of the preceding paragraphs in terms of a graph identical to Figure 7 of Chapter 6. The indifference curves of that graph become isoquants, and the budget line represents a fixed level of cost. The slope of the budget line will depend on the relative prices of resources. Output is maximized for a fixed level of expenditure at the tangency of an isoquant and the budget line, and at this point the equimarginal principle holds because the slopes of the two lines are equal.)

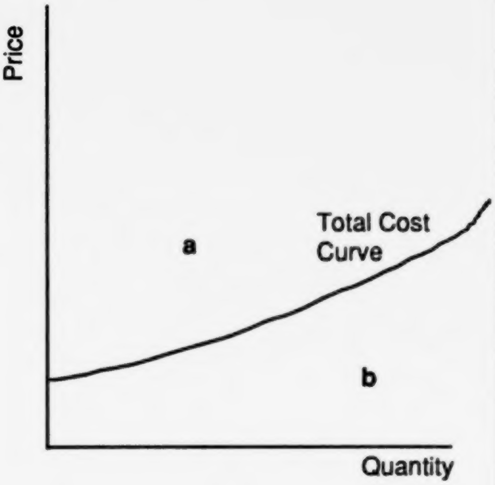


Figure 10.1 The region below the total cost curve is not attainable.

Once a minimum total cost curve is determined, the marginal cost curve can be found from it. Marginal cost is the additional cost of producing one more unit of output, or the change in cost divided by the change in output. This should not be confused with average cost, which is total cost divided by total output. Marginal cost plays an important role in the economist's story of the firm; average cost plays a secondary role.

At the risk of repeating material that is becoming stale if you fully understand the logic of the maximization principle, Table 2 illustrates the logic of finding the level of output at which profit is maximized. Columns 1 and 2 show the total cost curve, and columns 1 and 3 show the demand curve. The total

TABLE 2
COMPUTATION OF THE BEST OUTPUT FOR THE FIRM

(1) OUTPUT	(2) TOTAL COST	(3) PRICE	(4) TOTAL REVENUE	(5) MARGINAL REVENUE	(6) MARGINAL COST
0	10	—	—	—	—
1	20	18	18	18	10
2	31	17	34	16	11
3	43	16	48	14	12
4	56	15	60	12	13
5	70	14	70	10	14
6	85	13	78	8	15

revenue curve comes from the demand curve; it is price multiplied by output. If the firm can sell four units at $15 each, its total revenue will be $60.

Since marginal revenue is the change in total revenue when another unit is produced and sold, to compute column five we simply look at the total revenue column and compute how much it changes when output changes. When output increases from 2 to 3 in Table 2, total revenue increases from 34 to 48, or by 14. This 14 is the marginal revenue of the third unit of output. Similarly, since marginal cost is the change in total cost from adding another unit of output, to obtain column 6 we simply look at the total cost column and ask by how much it changes when output changes. When output increases from 2 to 3 in Table 2, total cost increases by 12, from 31 to 43. This 12 is the marginal cost of the third unit of output.

Producing the first unit of output in Table 2 is worthwhile because it adds 18 to revenues and 10 to costs. Paying 10 to get 18 is a good deal, and the firm should do it. For the same reason, the second and third units are worth producing. To produce the fourth unit, the firm must pay 13 to get 12. This is not a good deal, so the level of output that maximizes profits is 3. (If it is possible to pick fractional units, it would be worthwhile to go just a bit beyond 3, to the point where marginal revenue exactly equals marginal cost.)

Notice that there is a cost in Column 2 when the firm produces zero output. This is a **fixed** cost, a cost that exists when nothing is produced. For a real-world firm fixed cost would include rental payments, insurance payments, and depreciation. If fixed costs are high enough, there may be no level of output for which a profit exists. In this case the solution which the maximization principle gives minimizes losses. If losses get big enough, and become permanent, the firm may find the best way to minimize losses is to shut down completely and cease to exist as a firm.

The condition that the firm must set marginal revenue to marginal cost can be expressed in another way. Marginal cost is equal to the change in cost divided by the change in output. Since the equimarginal principle must hold if the firm is on (not above) its total cost curve, the change in cost can come from a change in any of the resources. (This will be exactly true in the world of calculus, where changes are so small that they approach zero. It will only be approximately true if changes are larger.) Thus marginal cost is equal to marginal resource cost divided by marginal product, a change in cost divided by a change in output. Hence the condition for maximizing profits is:

MR = MRC/MP.

In this equation we can see the role of all three constraints which the last chapter discussed.

THE PROFIT-MAXIMIZING LEVEL OF AN INPUT

The determination of the profit-maximizing level of an input is, like the determination of the profit-maximizing level of output, a mathematical problem if there is perfect knowledge of the supply curves of resources, the production function, and the demand curve. It is another application of the maximization principle, which says that the best level of an input is that level for which its marginal benefit to the firm, the extra money the firm can obtain by hiring or buying the input, just equals the marginal cost to the firm of hiring or buying the input. In the jargon of economists, the **marginal revenue product** of an input should equal the marginal resource cost.

The previous chapter showed that marginal resource cost depends only on the supply curve of the input. The marginal revenue product is the extra revenue which the firm can obtain from hiring another unit of the resource. It depends both on the extra output which the input produces and on the extra revenue the firm can obtain from each extra unit produced. For example, if adding another unit of an input can increase output by three, and selling an extra unit of output increases revenues by $4.00, the marginal revenue product is $12.00. If it costs an extra $10.00 to hire or buy the input, the firm will increase profit by $2.00 if it does buy or hire it. If the input costs an extra $14.00, the firm will decrease profits by $2.00 by using it, and should consider a reduction in the level of this particular input.

Because marginal revenue product equals the marginal product of a resource multiplied by the marginal revenue of the output, the condition that marginal resource cost should equal marginal revenue product, or:

MRC = MRP

can be written as

MRC = MP x MR = Marginal Product x Marginal Revenue.

This formula can, with a bit of algebraic manipulation, be turned into the condition for profit-maximizing output, or:

MR = MC = MRC/MP

TABLE 3
COMPUTATION OF THE BEST INPUT LEVEL FOR THE FIRM

1 Marginal Revenue Product	2 Marginal Resource Cost	3 Labor	4 Output	5 Marginal Product	6 Marginal Revenue	7 Marginal Cost
—	—	0	0	—	—	—
$27.30	$10.00	1	13	13	$2.10	$ .77
14.70	10.00	2	20	7	2.10	1.43
12.60	10.00	3	26	6	2.10	1.67
10.50	10.00	4	31	5	2.10	2.00
8.40	10.00	5	35	4	2.10	2.50

A simple numerical example shows the equivalence of the two approaches—by finding optimal output or optimal input—to maximizing profits and the role of the three constraints. Assume that a firm is a price taker both as a seller of output and as a buyer of inputs, and that it can sell as much as it wants at a price of $2.10 per unit and buy as much labor as it wants for $10.00 per unit. Assume also that the amount of capital is fixed at two, that capital and labor are the only inputs, and that the firm faces the production function shown earlier in Table 1.

Columns 3 and 4 of Table 3 show data taken from the production function. Column 5 shows the marginal product of labor, derived from columns 3 and 4. Column 2 shows the marginal resource cost of labor, which comes from the assumption that the firm can buy as much labor as it wants at $10.00 per unit. Column 6 shows marginal revenue, which comes from the assumption that the firm can sell as much as it wants at $2.10 per unit of output.

To find the profit-maximizing amount of labor, the firm must compare the extra cost of another unit of labor with the extra revenue that the extra labor adds. The extra cost is the marginal resource cost, shown in column 2. The extra revenue is the marginal revenue product, the value to the firm of the extra output which the additional labor produces. Column 1 shows marginal revenue product, which is the product of marginal revenue and marginal product (columns 5 and 6). By comparing the marginal revenue product to the marginal resource cost, one can immediately see that the profit-maximizing amount of labor is four. Hiring the fourth unit of labor adds $10.50 to revenue and $10.00 to cost, so profits increase. Hiring a fifth unit adds less to revenue than to costs, $8.40 and $10.00 respectively, so it is not a profitable unit to hire. If the firm hires four units of labor, the production function says that it will produce 31 units of output.

The previous section explained that to calculate profit-maximizing output, one needed to compare marginal cost of output and marginal revenue. Column 6 of Table 3 contains the marginal revenue. To find the marginal cost of output, one must compute what it costs to produce another unit. When the firm hires the first unit of labor, it adds 13 to output and $10.00 to cost. However, we do not want to know what an extra 13 cost, but what an extra *one* costs. To find it, we divide the $10.00 into 13 equal parts, and call the result the extra cost of producing one more. In other words, to get the marginal cost of output in column 7 of Table 3, we need to divide the marginal resource cost by the marginal product.

Comparing columns 6 and 7, one can immediately see that 31 is the profit-maximizing level of output. Producing more or less will reduce profits. If the firm produces more than 31, it adds $2.00 in revenue for each extra unit, but it also adds $2.50 in cost. If it produces less, it cuts costs, but not by as much as it cuts revenues.

Checking the production function at 31 units of output, one sees that one needs four units of labor. Hence finding profit-maximizing output and profit-maximizing input are two different ways to arrange the information from the production function, the supply of resources, and the demand for output. Both ways give identical results.

PROFIT MAXIMIZING IN REAL FIRMS

A firm maximizes profits by setting marginal cost equal to marginal benefit. However, when one examines real-world firms, one finds few managers who make decisions by trying to determine marginal costs and marginal benefits, and then comparing them. In the 1960s *Business Week* found a case in which a firm did use marginal analysis, and decided that it was unusual enough to be newsworthy. Continental Airlines made decisions about which flights to fly by comparing the extra costs of a flight with the revenues it would generate. At the time most airlines made these decisions on the basis of the percentage of seats filled on a flight.

Some economists have tried to construct a theory of the firm in which the firm decides prices by a mark-up over costs. The attraction of this sort of theory is that most sellers, when asked to explain how they determine the prices they charge, talk in terms of mark-ups (which they sometimes call "profit margins"). The problem with mark-up theories is that they have difficulty in explaining the percentage size of the mark-up (when they bother trying to explain it at all). Grocery stores, for example, mark up different products by different percentages, and they have a much smaller average mark-up than furniture stores have.

(Mathematically one could determine the optimal mark-up for each product. However, this will result in maximizing profits and will give a solution identical to the marginal-cost marginal-benefit solution. It is subject to the same criticism the marginal-cost marginal-benefit solution is subject to: it does not seem to describe the way real-world firms actually set prices. See Question 3.)

Most economists see mark-up pricing as a rule-of-thumb way which businesses conduct their affairs. Firms usually do not have the information needed to compute marginal costs and revenues. Rather they find rules or guidelines which work, and stick with them as long as they perform satisfactorily. If a firm marks up a product by 50% and finds that it does not make a profit at that price, it tries another percentage. When it finally finds a mark-up which generates a profit, it will stick with it. Real businesses use much more trial and error than sophisticated, mathematical analysis.

Even if firms do not attempt to set marginal cost equal to marginal benefit, (and even if they do not attempt to maximize profits), they may end up doing so. In an industry in which firms can enter easily, firms with production and pricing decisions far from those needed for profit maximization probably will not survive long.

There is an analogy here to the theory of evolution. Animals and plants which best find food sources and which best avoid predators will survive and reproduce. As a result one can view the development of a species as a competition among genes. A gene which gives the organism which contains it access to new food sources, or better ways to exploit old food sources, or better ways to avoid predation is more likely to be passed on to the next generation, and thus survive, than those genes which are competing with it in the gene pool. It is not the *intention* of genes to do these things, but the result is that genes act *as if* they maximize, and biologists have found that analysis assuming that genes maximize their chances for survival can give insights. Similarly, in a competition for profit, those firms which do maximize, whether intentionally or by accident, stand the best chance of survival and growth. As a result, firms may end up close to decisions on output and price which they would make if they knew their marginal costs and marginal benefits and made decisions based on them.

Finally, you should realize that a firm which explicitly sets out with a goal of maximizing profits may actually do less well than one which takes an indirect approach to this goal. Setting up a goal such as "quality at a reasonable price" may be better for both employee morale and consumer reception than a more straightforward "profit-first" goal.

The result that firms maximize profits, or come reasonably close to it, is important when we ask how well firms maximizing profits serve the public's interests. This will be the topic of Chapters 11 and 14. It also has some predictions about the effects of government restrictions and controls, one of which the next section considers.

AN APPLICATION OF THE THEORY OF PROFIT MAXIMIZATION

The theory of the firm says that the firm produces at that level of output for which marginal revenue equals marginal cost. Thus given the data in Table 4, the firm can be

TABLE 4
DATA FOR HYPOTHETICAL FIRM

OUTPUT	MC	MR	TOTAL COST	PRICE
1	$4.5	$16	$4.5	$16
2	4.6	14	9.1	15
3	4.7	12	13.8	14
4	4.8	10	18.6	13
5	4.9	8	23.5	12
6	5.0	6	28.5	11
7	5.1	4	33.6	10

expected to produce the sixth unit but not the seventh. The sixth unit *adds* 6 to revenues and only 5 to costs, so profit increases by 1. The seventh unit adds more to costs than it adds to revenue, so it decreases profit.

Suppose that the government decides to try to lower the price in this market by putting price controls on one of the key inputs. Would these controls result in lower prices for consumers? It may be a bit surprising, but the theory of the firm suggests that price controls on an input which is sold by price takers will not lower the price of the final product but raise it.

Government price controls on an input affect a supply curve of a resource. (Here is where the assumption that the input is sold by price takers is important. If it is sold by a monopolist, there will be no supply curve, and no prediction can be made about what effects price controls on the input will have. See question 5.) The industry initially will

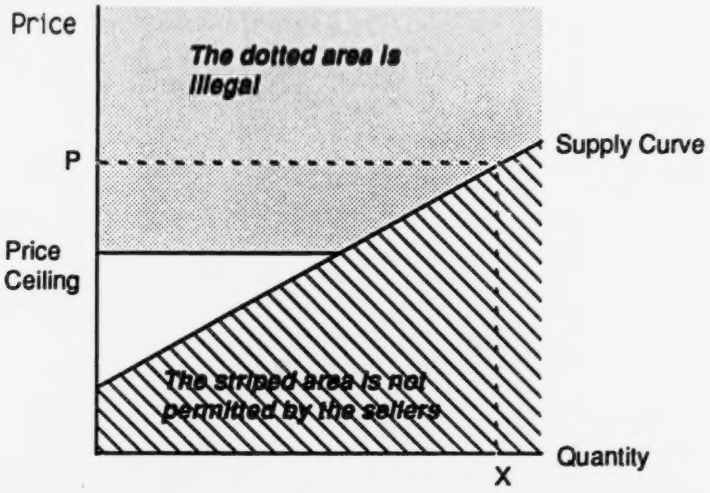

Price
The dotted area is illegal
P
Supply Curve
Price Ceiling
The striped area is not permitted by the sellers
Quantity
X

Figure 10.2 With a price ceiling only the blank region is available to resource buyers.

be using some amount of this resource, say X, and paying a price, say P. The controls outlaw price P and make only lower prices legal. Will the amount X still be available at the new, lower price? Recall that one can view a supply curve as a boundary, a boundary which limits buyers. The combination of amount X and the lower price is in the "not-permitted" area which the supply curve defines. At the new, lower price less than X of the resource will be available. This is a key link in the argument that follows.

The price of the resource will decline, and so may average costs. But *marginal* costs will not fall at the level of output the firm produces because the supply of the resource is cut. The representative firm will use as much of this cheaper resource as it can, but it will no longer be able to get as much of the resource as it got before. Once it cannot get more, it will have to find substitutes which will cost at least as much per unit of output (or else the firm would have preferred to use these substitutes in the first place). The increase in marginal costs, which the reduced availability of the input causes, will lead to higher prices for the consumer, though lower average costs may lead to higher profits for the firm.

In terms of Table 4, the effect of price controls on the input is to lower the marginal costs for lower levels of output, but to raise them for the levels of output at which the firm normally produces. New numbers in the MC column of the table might be: 4.1, 4.2, 4.3, 4.4, 6.1, 6.3, 6.5. These numbers change the output level which is most profitable. You can check for yourself to see if profits are increased or decreased, and whether price is lower or higher.

When the wage-price controls which President Nixon imposed late in his first term were removed, the controls on domestically produced oil were kept. The U.S. government could not control the price of crude oil which foreigners produced and, as a result, refiners which were heavily dependent on foreign crude oil would have been at a serious disadvantage without help. The government imposed further provisions that refiners which controlled large amounts of domestic production pay refiners which did not. Though this re-allocation complicates the computation of costs, the cost of the foreign oil was the approximate marginal cost of crude oil.

The issue of oil decontrol bubbled up throughout the Ford and Carter administrations. The Carter administration, in particular, opposed oil decontrol, arguing that it would raise gasoline prices. Domestic oil was finally decontrolled early in the Reagan administration.

The argument of the Carter administration, that price controls on crude oil kept the price of gasoline low, had almost unanimous support in the news media and wide public appeal. But the argument is difficult to present in a logical, theoretical manner. It apparently rested on a belief that price is based on average costs plus some markup. Sellers often use mark-up rules for pricing, but, as the last section mentioned, a problem with mark-up theories is that they do not adequately explain the percentage sellers use. In contrast, the economic theory of the firm suggests that price controls on domestic crude oil would not decrease the prices consumers paid for gasoline. The major effect of the controls should be, according to this theory, to transfer profits from the domestic producers to refiners, and to strengthen the OPEC oil cartel by discouraging domestic production.

Some evidence supports the hypothesis that the effect of price controls on crude oil was to transfer profits among firms. During the debate about whether to continue or eliminate crude-oil price controls, those large oil companies which were primarily refiners tended to favor continuing controls while those which were relatively large producers tended to favor abolishing controls. In at least one company the refining division initially lobbied to keep controls and the production division to abolish them; when the highest levels of management became aware of the difference, one division was stopped. Mobil Oil, which had been running a series of public-service ads proclaiming the virtues of the free market, was primarily a refiner, and came out in favor of controls. This division of the oil industry is not surprising if a major effect of the controls on crude oils was to shift profits from producers to refiners.

When price controls on crude oil were finally eliminated, news media reported many predictions that gasoline prices would rapidly increase. There was no immediate change that could be attributed to the elimination, and within a year the price of gasoline had fallen. These results suggest (but of course cannot prove) that the economic theory of the firm gave a better explanation of the effects of government policy than did the various markup theories which the government and the news media preferred.

SUMMARY

This chapter showed how the maximization principle can be applied to the firm. It explored the assumption that firms maximize profits, pointing out some of the ambiguities of this assumption. It then explained the rule that marginal cost equal marginal revenue, and the rule that marginal resource cost equal marginal revenue product. It showed that these two rules led to the same results but represented different ways of using the information from the three constraints which a firm faces, the production function, the demand curve, and supply curves of resources.

QUESTIONS

1. (a) Suppose there are 30 people in a room with the average age of 40 years. A fifteen year old boy walks in and the average age of the group becomes 40.5. Is this possible?

(b) A firm is producing 30 units of output at an average cost of $40. Adding another unit will increase costs by $15 (this is the marginal cost) and the average cost will rise to $40.5. Is this possible?

(c) If average cost is falling, what must be true of marginal cost?

2. The Ajax Gizmo Company can sell gizmos for $.50 each. Based on the table below, how much of resource X should it purchase? Explain your answer.

Amount of X	Marginal Product	Price Necessary to Buy Amount Listed	Marginal Resource Cost
0			
	10		
1		$2.00	
	9		
2		2.20	
	8		
3		2.40	
	7		
4		2.60	
	6		
5		2.80	
	5		
6		3.00	

3. a) A firm has constant average costs of $10. The elasticity of demand is also constant, at 2. What is the profit-maximizing markup? (Hint: This problem looks impossible, but equation 9 in Chapter 5 makes it quite simple.)

b) A firm finds that producing one hundred units of output and selling them for $200 each maximizes profits. It also has found that to sell the 101st unit, it must lower price by $1. What must its marginal costs be?

4. Table 3 in the text illustrates two different ways to compute the profit-maximizing position for a firm. The profit-maximizing firm will hire four units of labor and produce 31 units of output.

a) Suppose that there are fixed costs of $15.00 that must be paid to capital. (Recall that capital is held fixed at two units.) What will the profit of the firm be? (Hint: Compute the total revenue and total cost columns in the table below, and then subtract the latter from the former to find profit.)

b) What will happen to the profit-maximizing output if fixed costs are decreased to 10? If they are eliminated? (Hint: What will a change in fixed costs do to the marginals in Table 3?)

c) Suppose that the workers run the firm, and that they want to maximize profit per worker, not total profit. How much will they produce and how many workers will the firm hire assuming that fixed costs are $15? (Hint: Compute the profit-per-worker column in the table below.)

Labor	Output	Total Revenue	Total Cost	Profit	Profit per Worker
0	0	0	$15	-$15	—
1	13				
2	20				
3	26				
4	31				
5	35				

(Comment: Yugoslavia has tried to implement a system of worker management. One problem which has plagued them is high unemployment.)

5. Table 2 of this chapter shows a seller maximizing profit at three units of output. Suppose the government puts a price ceiling of $15 on this seller.

a) Recompute the total revenue and marginal revenue columns of Table 2. (Hint: What is the highest price the seller can charge if it sells only one unit? If it sells six units? Be careful.)

b) What is the profit-maximizing level of output with the price ceiling?

c) Compare your answer to what Chapter 6 says are the effects of a price ceiling on output. What do you conclude? (Hint: Check Chapter 4 for assumptions on the number of buyers and sellers needed for supply and demand curves.)

6. (a) Suppose a retailer pays $500 for a product and sells each for $600, a markup which covers costs and earns him a normal profit. The manufacturer of the product then drops the price that the retailer must pay to $250. Should the retailer keep the price at $600 for the product in stock? Explain.

(b) Suppose the price change had gone the other way. The retailer initially bought the product for $250 and sold it for $350. Then the manufacturer raised the price to $500. Should the retailer continue selling the product on hand at $350, and only begin to charge a higher price when he starts selling the items for which he paid $500? Explain.

(c) If a seller is concerned about the possibility of a price cut, how can he protect himself? Explain.

(d) Regulations that would prohibit retailers from repricing products have been suggested as a way to help consumers in times of inflation. Such regulations would not allow a retailer to raise the prices of items which had been already priced (though he could lower them). Thus if the price of reordering increased, the retailer could not increase the prices of items already on the shelves. What sort of change in behavior on the part of sellers would such regulation encourage? Would buyers necessarily be better off?

7. Below are two tables similar to Table 3 in the chapter but with different numbers. (a) Complete the tables and find the profit-maximizing levels of output and labor.

1 Marginal Revenue Product	2 Marginal Resource Cost	3 Labor	4 Output	5 Marginal Product	6 Marginal Revenue	7 Marginal Cost
$30.00	$9.00	1		10		
25.00	10.00	2		10		
20.00	11.00	3		10		
15.00	12.00	4		10		
10.00	13.00	5		10		

1 Marginal Revenue Product	2 Marginal Resource Cost	3 Labor	4 Output	5 Marginal Product	6 Marginal Revenue	7 Marginal Cost
	3.00	.5	1		4.00	
	3.00	1.0	2		3.50	
	3.00	1.6	3		3.00	
	3.00	2.3	4		2.50	
	3.00	3.1	5		2.00	

(b) Compute the supply curve of labor that exists for the first table, and the demand curve for output that exists for the second table.

Amount of labor	Wage	Amount of Output	Price
1	1		
2	2		
3	3		
4	4		
5	5		

8. Answer the next seven questions on the information in this table:

SUMMARY OF ACE GIZMO COMPANY

Labor Hired: 5
Cost per unit of labor: $5000
Capital hired: 5
Cost per unit of capital: $10000
Total Production: 81
Price Charged: $1020
Sales: 81
Total Revenue: ________
Total Cost: $75000
Profit: ________

If another unit of labor is hired, output will increase by about 6.5 Gizmos.

If another unit of capital is hired, output will increase by about 10.6 Gizmos.

A 1% change in price will cause a 10% change in sales.

First, some easy questions.

a) What will the total revenue and profit of the company be?

b) A consultant estimates that if the firm hires 10 units of labor and 10 units of capital, Ace Gizmo will produce 168 gizmos. The president of Ace Gizmo questions this estimate, arguing that it cannot be correct because it violates the law of diminishing returns. Evaluate the president's argument.

c) When it produces 81 gizmos, the firm finds that it can sell them all at a price of $1020. If it changes price by 1%, it changes sales by 10%. Is Ace Gizmo Company a price taker or a price searcher?

d) What is the price elasticity of demand?

And now some questions that are a bit more challenging.

e) Is the firm producing the 81 gizmos as cheaply as possible? Explain how you can tell, and what changes, if any, would increase output without increasing cost.

f) What is the marginal revenue of the 81st gizmo?

g) Estimate the marginal revenue product of the fifth unit of labor. Is this bigger or smaller than its cost?

9. A housewife is buying four 12 ounce cans of peas for $.60 each and three 16 ounce cans for $.72 each. Using the equimarginal priniciple, determine if she is minimizing her costs of buying peas. If she is not, explain how she can shift her money to increase the amount of peas without spending more.

CHAPTER 11

EFFICIENCY

After you read this chapter and work the problems at the end, you should be able to:

Define economic efficiency, Pareto optimality, welfare loss.

Explain the three situations which allow an increase in economic efficiency.

List at least two assumptions needed before one can take economic efficiency seriously as a normative criterion.

Explain why a competitive market economy can be economically efficient.

Chapter 3 explored the relationship between an individual's interests and the interests of the group. It found that sometimes there was conflict and sometimes there was harmony between self-interest and group interest. In the seven chapters since, we have developed enough analytical tools to come back for a more sophisticated look at this question of the individual versus the group. In the next five chapters we ask what sort of economic system, if any, brings harmony rather than conflict between the individual and the group. For economists, the concept of economic efficiency is a key to discussing this question.

DEFINITIONS

Efficiency is a relative term. It is vital that this point be understood. Efficiency is never absolute; it is always relative to some criterion. This can be seen when one asks if farms are more efficient in the United States or in China. The farming techniques in China are more efficient than those in the United States when measured in terms of output per unit of land, or output per unit of fossil fuel, or output per unit of machinery. The farms in the United States are far more efficient in terms of output per man hour. The statement that farms in one country are more efficient than farms in another makes no sense unless the criterion on which efficiency is measured is given.

Economic efficiency, which economists use when they discuss efficiency, is measured in terms of value. A change which increases value is an efficient change and any change which decreases value is an inefficient change. A situation which is economically efficient may be inefficient when judged on different criteria. A story may make this concept clear. In the late 1960s an electric utility in New York wanted to build a storage facility at a location called Storm King Mountain. The plans called for electrical pumps that would force water up the mountain during the night to a storage reservoir. During the day the water would flow back down the mountain to produce electricity. However, the laws of physics reveal that it takes more electricity to pump the water up the mountain than can be recovered when it flows back down. Thus this project cannot be energy efficient. Still the company wanted to build it because they believed that it was efficient economically.

To understand the company's point of view, a brief digression into methods of generating electricity is necessary. Most electrical utilities rely primarily on very large power plants that are fueled either by coal or nuclear fuel. These plants take a long time to heat up and a long time to slow down; they run best when they run at a constant rate. But the demand for electricity varies over the course of a day. At 2:00 AM, for example, it is very low. At 4:00 PM it is quite high. The companies have a number of options to meet this daily fluctuation. On one extreme, they could run main power plants at a high daily level, or peak demand, and simply let any surplus electricity go to waste. At the other extreme, they could run the main plants at the lowest daily level and also run "peaking plants" to provide additional power when it is needed. A peaking plant is a generator, often fueled by natural gas or diesel fuel, which can be turned on and off quickly. Though much cheaper to build, its fuel cost per kilowatt are higher than with a main plant.

The utility company in the Storm King case would have used electricity generated during periods of slack demand, when electricity was "cheap," to pump water, and would have generated electricity with the water during periods of peak demand, when electricity was "dear." The company believed that this was the cheapest way it could meet the daily demand for electricity, and hence thought that it was economically efficient.

Value is subjective. A thing has value only if someone wants it. How then can we know if value is maximized? If there is some change which makes someone feel better off, but making this change does not make anyone feel worse off, then the original situation was not one of highest value. Improvement was possible. When the highest value is reached, then any possible change that helps anyone must harm someone else. This way of defining economic efficiency is called **Pareto optimality**. It is named after Vilfredo Pareto, an early mathematical economist. He was an Italian who wrote in the late 19th century.

Economist are interested in economic efficiency for two reasons, one positive and the other normative. The positive reason is based on the observation that people search for value. We see this search for value vividly illustrated in the occupations of pimp, drug pusher, and hit man; given enough money, any occupation, no matter how immoral or risky, will attract people. On the theoretical level, we have seen this search for value in discussing utility maximization and profit maximization. The search for value is the driving force of market (and perhaps most nonmarket) economies. If there are situations in which there is unexploited value, that is, value which is possible but which no one obtains, the economist needs to explain why someone does not find a way to capture this value.

The normative reason stems from a desire to make policy recommendations. It is possible to discuss some aspects of policy without normative assumptions. An economist can predict, for example, whether a policy will or will not achieve the goals set for it. But economists often want to do more. They often want to compare two policies or two situations and decide which is better. To decide which is better requires some sort of basis for ranking situations. Thus if they want to ask whether government regulation of utility prices, or a tariff on steel, or a program to train unskilled workers helps society, economists need a criterion on which to base their answer. Economists generally use the criterion of economic efficiency to evaluate situations, though they often supplement it with other considerations because economic efficiency is not the only way to judge the relative merits of two situations.

The value maximized in the notion of economic efficiency reflects the goals people have. The concept of economic efficiency treats all goals as equally valid; no goals are considered better than other goals (with the one exception noted below). Not everyone agrees. Judging goals has been a central feature of the Judeo-Christian tradition. Generally this tradition has condemned as immoral goal-seeking that emphasizes the most narrow individualism such as hedonism. To be moral, people must take into consideration as a goal the well-being of some others, including family or clan members and others who are members of a community grouping.

The one goal which economists outlaw is the desire to harm others, or envy. If one person feels unhappy whenever another feels better off, there is no possible way to rank situations. In addition, envy is destructive of cooperation, and so it has few supporters.

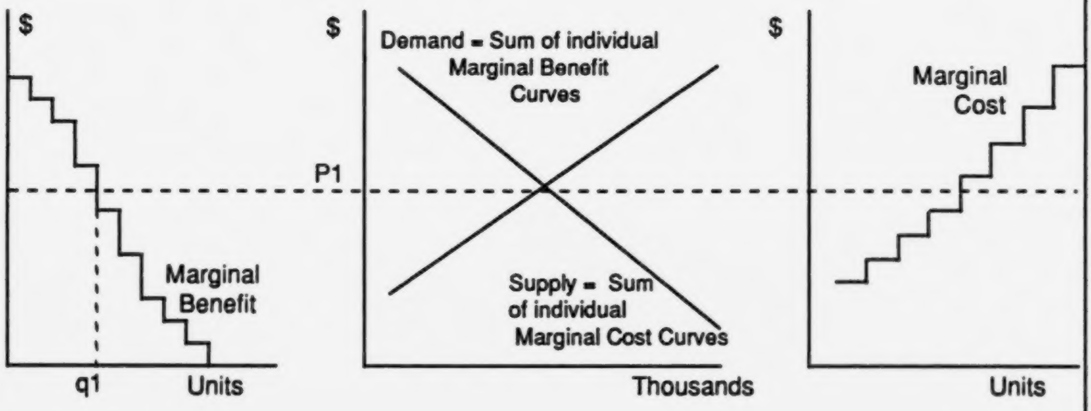
Not all goals are equal in determining value. The goals of some are given more emphasis than the goals of others. In a market economy, the goals of the rich are given more weight than the goals of the poor. The rich have more dollar votes. If one dislikes the goals that people pursue, or if one believes that the goals of some people—the rich--are given too much emphasis and the goals of others—the poor—are given too little, one may believe that an economically efficient situation is inferior to one that is economically inefficient.

EFFICIENCY AND MARKETS

Adam Smith observed that people pursuing their own interests could, if guided by a market, serve the public interest. Questions about this observation—do the people pursuing self-interest serve the public interest as well as they could, and if conditions are changed, will they serve the public interest better or worse—have led economists to explore the efficiency of the model of an exchange economy.

Previous chapters have developed the analytical tools needed to show that a simple model of an exchange economy can be eco-

Figure 11.1 A buyer, a seller, and the market.



Representative Buyer

The Market

Representative Seller

nomically efficient. In this model, all transactors must be price takers. This assumption means that supply and demand curves can be used to represent all markets. The supply and demand curves of a representative market, the bread market, are shown in the center of Figure 1. The graph on the left of Figure 1 shows the viewpoint of the typical buyer, Jane Doe, who is one buyer of a great many buyers. On the right is a graph showing the viewpoint of the typical producer, Martha Smith. The price that the market produces will be P1, the market clearing price.

Jane Doe's demand curve for bread is downward sloping. At a high price, she buys little, and at a low price she buys much. The reason for this behavior is that, as she gets more bread, the value of another loaf declines. In fact, one can go further and argue that her demand curve shows the marginal benefit of bread to her. If she is willing to buy five loaves for $.50 a loaf, but not the sixth loaf at this price, then the value of the sixth loaf, or its marginal benefit, must be less than $.50. If the price must drop to $.45 per loaf before she buys the sixth loaf, her behavior reveals that the marginal benefit of the sixth loaf is $.45. Her willingness to pay, revealed in her demand curve, measures her marginal benefit from bread.

Jane Doe, as a price taker, can buy all she wants at the market price. To her the market price appears to be a supply curve. It is also her marginal cost of buying bread because it shows how much an additional loaf will cost. Assuming that she maximizes utility, she will buy until marginal benefit equals marginal cost, or to the point at which the price line crosses her demand curve. At price, P1 in Figure 1, she buys q1 loaves.

Jane Doe is only one of a great many other buyers. To find out how much all buyers will take at each price, one needs to add up how much each individual buys at each price. (This is not a problem in an abstract model, but is clearly impossible in the real world.) The summation of individual demand curves yields the market demand curve. Because it is found by adding marginal benefit curves, the market demand curve shows the marginal benefit to buyers. This identity of demand curves and marginal benefit curves is vital in considering the efficiency of a model of a market economy.

Marginal costs and benefits look different from the point of view of Martha Smith, a typical seller in this market. To her the fixed price shows the marginal benefits of a transaction. The price tells her how much more revenue she gets from selling another loaf of bread. Her marginal cost curve reflects the value of the resources she must add to make another loaf of bread. These costs depend on two things. First, they depend on how much extra resources are needed to produce another loaf of bread, or the marginal productivity of resources. This productivity is determined by the technology of making bread. Second, they depend on the prices of the needed extra resources, which will depend on the alternative uses of these resources. If the resources needed to make bread are high priced, these resources must, in our simple model, have alternative uses which are highly valued. Marginal cost curves can slope upward or downward (recall returns to scale), but only an upward sloping curve will give us sellers who are price takers.

Assuming that Martha Smith wants to maximize profits, she will find that level of output for which marginal revenue, which is price, equals marginal cost. This profit maximizing output can be found on the graph by finding the intersection of the price line and the seller's marginal cost curve. For any price, the marginal cost curve tells how many loaves the profit-maximizing seller will produce. Since a supply curve also tells how much a seller will sell at given prices, the marginal cost curve must be a supply curve.

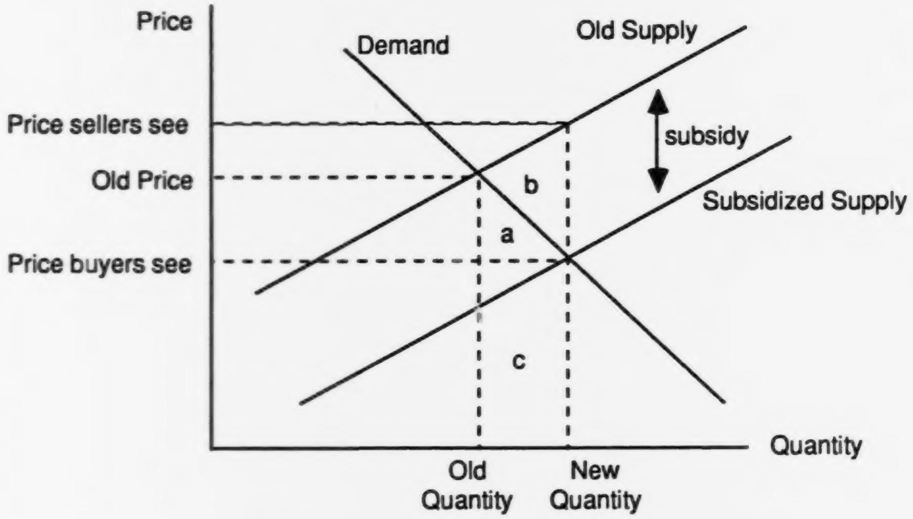
The individual seller is only one of a great many sellers. The market supply curve is obtained by seeing what each seller does at a price, then adding up all the outputs at that price. (Again, this presents no problems in an abstract model, but is clearly impossible in the real world.) The result will be a supply curve that can also be interpreted as a marginal cost curve. (A technical note: to get a market supply curve by summing up individual supply curves [or a market demand curve by summing up individual demand curves], we need to assume that a change in price or quantity will not affect any of the factors held fixed. Thus if higher prices and quantities for output bid up the price of resources, the individual supply curves would change, and the assumption that other factors are held fixed is violated. In this case a simple addition would not be correct. This qualification is ignored in the discussion below.)

The market supply and demand curves in Figure 1 are identical to those discussed in Chapter 4. The point of the discussion above (you are normal if you are asking why bother with all this abstract reasoning) is that the demand curve represents a marginal-benefit-to-buyers curve and the supply curve represents a marginal-cost-to-sellers curve.

However, we are not done yet. In the model of the exchange economy one can (given a few hidden assumptions) also show that in the goods market the marginal cost to the sellers is the same as the marginal cost to the buyers. Marginal costs to sellers are determined by what the firm must pay for resources necessary to produce one more unit of output. To produce one more unit of commodity A, a seller must bid resources away from another use. If in the alternative use the resources were producing output worth an added six dollars, the seller could not bid them away for a mere five. If those resources add six dollars worth of output somewhere else, the seller of product A will have to pay at least six dollars for them. But he will not have to pay very much more for them. If we assume perfect information and the absence of transactions costs, then any amount, however small, over six dollars will cause resources to move to product A.

The demand curve represents marginal benefit to buyers, and, because sellers take the price from it, to sellers as well. The supply curve represents marginal cost not only to sellers, but also to buyers. It tells both buyers and sellers what the value of the resources needed to produce the product is. The maximization principle says that to maximize net benefit, marginal benefit must equal marginal cost. As a result, the optimal output, the one which maximizes value to society, will be the output at which supply and demand intersect, the amount which the market provides. This is a remarkable conclusion: Adam Smith was correct in saying that pur-

Figure 11.2 A bread subsidy.



suit of self-interest could lead to a socially desirable result. A look at two other possibilities may help the reader to grasp the meaning of this conclusion.

Suppose that the government decides to help consumers by lowering the price of bread with a subsidy to sellers. Figure 2 indicates the effects of this policy. The subsidy makes the price of bread look higher to sellers, and they produce more. It makes the price look lower to buyers, and they buy more. Since both buyers and sellers would appear to be happy with this result, can we conclude that this subsidy helps the economy? The answer is a surprising "No."

The key to seeing why the subsidy does not increase value is to realize what scarcity implies. The production of more bread requires more resources, and these resources must be drawn from other uses. To get the extra bread, consumers must do with less of other goods.

In the simple model of Figure 2, the value and cost of the extra bread can be measured. The value of the extra bread is the amount which people are willing to pay for it. Consumers were not willing to buy the extra bread at the old price, so it must be less valuable to them than the old price. It will, in fact, be equal to the extra area under the demand curve, or area a+b in Figure 2. But production of this extra bread requires resources that have more value than the value of the extra bread. Marginal cost is the value that must be paid to resources in order to produce one more loaf of bread, and with the subsidy this value exceeds the value of the loaf. The extra cost will be the area under the old supply curve, or area a+b+c. The subsidy causes businesses to take valuable resources and to transform them into less valuable output. There is a waste in this process, and the triangular area c is a measure of this waste. Economists call this area **welfare loss.**

(A technical note: Not all changes in behavior represent welfare loss. If a change makes John richer and Joe poorer, Joe will consume less but his lower welfare may be entirely offset by John's gain. Changes in behavior due to changes in income are ignored in discussing welfare loss. Only those

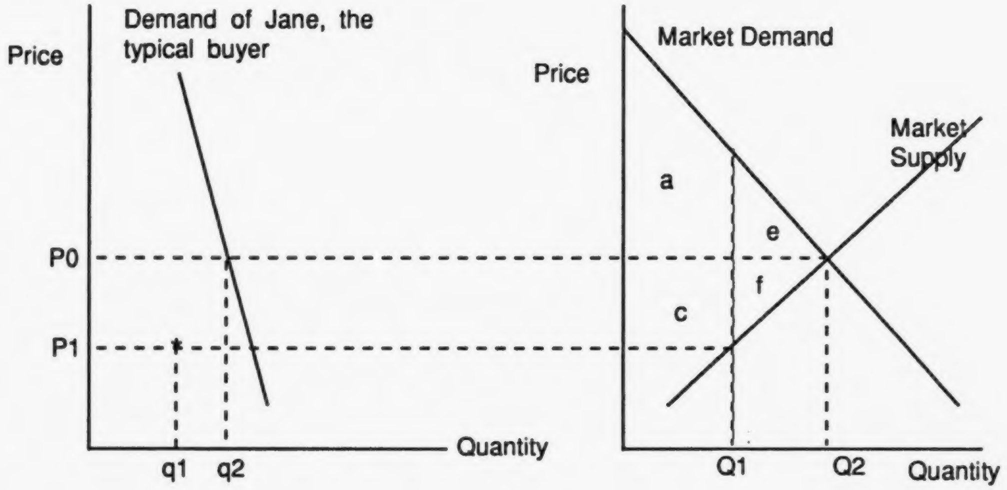


Figure 11.3 A restriction by a consumers' union.

changes that are substitution changes, where a person does more of one thing and less of another because of changes in relative prices are counted as a welfare loss. As a result, using regular demand curves to measure welfare loss may result in an error. Instead a special demand curve, one which captures only substitution effects, is the proper demand curve to use when drawing triangles of welfare loss. In many cases the difference will be very small, but for more information check an intermediate level textbook under the topic of "income-compensated demand curves.")

Though bread subsidies are wasteful, they can be politically popular. People see the good side of the subsidy, the cheap bread. What they usually do not understand is the bad side of the subsidy. They do not realize that the higher prices of other goods can be related to the low price of bread. An economic system is a complicated system, and many of the cause-effect relationships take a good deal of abstract thinking to comprehend.

If a government subsidy increasing the amount of bread does not help consumers, perhaps a contraction of the amount of bread will. The discussion of consumers' and producers' surpluses in Chapter 8 suggested that producers or consumers could gang up on the other and transfer surplus. Suppose bread buyers unite and restrict purchases of bread. This union, which changes our assumption of price taking, will forĉe down the price of bread.

Figure 3 illustrates the story of the consumers' union. For the union to be successful, each buyer must restrict purchases as Jane does in the the graph. Consumers as a whole will be better off as a result of this action because consumers' surplus will rise. Prior to the union, consumers surplus will be a+e in Figure 3, but after the price reduction, it will be a+c. As long as area *c* is greater than area *e*, the consumers' union benefits consumers.

However, the union does not benefit society as a whole because the Martha Smiths are as much members of society as are the Jane Does. The consumers are better off only because they have seized value that previously was already there, captured by the producers. But in the process of making this transfer, value equal to areas e+f was lost.

Furthermore, there are changes which could improve consumer well-being. If consumers could continue to buy Q1 at price P1, they would be willing to buy additional loaves at a higher price. Another loaf has greater value to them than the resources needed to make it. If they could keep the low price of P1 for quantity Q1, and then pay extra for additional loaves, they could capture some of the wasted value. By restricting bread production, resources have shifted from bread production to other uses which are less valuable to society than the value of foregone bread.

The unexploited value which the consumers' union creates would be tempting for each individual, and it is unlikely that such a union of many buyers could exist long because of the free rider effect discussed in Chapter 12. In practice the same result is obtained by requiring the farmer to sell grain to the government at a price below market equilibrium. The bread price is kept low by a system of rationing, either by queuing or coupon. Such systems are fairly common throughout the world whenever bread consumers have more political influence than peasants.

Discussion of the welfare loss caused by bread subsidies and by requirements that farmers sell their products to the government at prices below world market prices is not just of academic interest; such policies are widespread and ancient. Rome had bread subsidies almost 2000 years ago. Today various forms of bread subsidies and price ceilings exist in the Soviet Union, some of the countries of Eastern Europe, and parts of Africa. The price of bread can be kept so low that the government must pass laws prohibiting farmers from feeding their livestock bread rather than grain, a temptation only when the price of bread is less than the price of the grain used to make it.

This section has shown that there is some use of resources which maximizes value, and that it is fairly simple to construct a model in which markets yield this use. However, the reader should understand that with alternative assumptions other results can happen. Showing that markets in an imaginary world result in economic efficiency does not tell us that markets in the real world do, but only that they might. The conclusions of this section can be shown more elegantly and precisely using mathematical methods. But there is a cost to a mathematical presentation—it requires sophisticated mathematics. The discussion above, and questions 5 and 6 at the end of this chapter, try to give the same insights in a more intuitive manner.

PRODUCTION EFFICIENCY

There is another way to examine economic efficiency, and that is to look at what conditions are needed for an economy to be efficient. There are three conditions needed, and if any one of them is violated, there will

TABLE 1
PRODUCTION INEFFICIENCY

Marginal Product of Capital is	5 widgets or 6 getwids
Marginal Product of Labor is	5 widgets or 4 getwids

be unexploited value in the system.

One condition necessary for an economy to be economically efficient is that it be on its production-possibilities frontier. If it is *not* on the production-possibilities frontier, then with the given resources and technology more could be produced. Since greater production would increase value, any position below the production-possibilities frontier is inefficient. Notice that a great many points satisfy this condition of production efficiency—every point *on* the production-possibilities frontier is production efficient.

To be on the production-possibilities frontier, all resources must be used. Unemployed resources indicate that more goods and services could be produced, which would mean that the economy was not on the frontier initially. In addition, resources must be used properly. If society randomly assigns people to jobs or if it assigns jobs on the basis of political reliability, it will not produce as much as it could. It will require some people with little intellectual ability to perform jobs which require great intellectual ability, and it will require some people with little strength and endurance to perform jobs which demand much strength and endurance. If switching people among jobs can increase output, the original situation was not on the production-possibilities frontier and thus not economically efficient.

This requirement that resources must be used properly can be stated more technically. Production efficiency requires that an equimarginal principle be satisfied. It requires that the ratio of marginal products for any two resources be the same for all products. Table 1 illustrates a case in which this condition is not met. Here the ratio of the two marginal products for the production of widgets is (5/5) or 1 and the ratio of the two marginal products for getwids is (6/4) or 1 1/2.

To show that the situation in Table 1 is not production efficient, consider what happens if the two producers trade a unit of labor for a unit of capital, with the widget producer getting the labor and the getwid producer getting the capital. There will be no change in widgets produced because the loss in output from the lost unit of capital (5 widgets) is completely offset by the gain in output from the extra unit of labor (5 widgets). However, there will be more getwids. The extra unit of capital adds 6 getwids, while the loss of a unit of labor subtracts 4 getwids. There is a net gain of two getwids. Since the amount of production after the exchange of resources was more than the original amount, the economy could not have been on the production-possibilities frontier originally. Further, since more output has more value to consumers, the original use of resources was less efficient than the use of resources after the trade.

As a result of the trade of resources, marginal products should change. Since more capital is being used in producing getwids, its marginal product in getwid production should drop (by the law of diminishing returns). Since more labor is being used in producing widgets, its marginal product in widget production should drop. Hence some exchange of resources should bring the ratios of marginal products to equality.

TABLE 2
Exchange Inefficiency

	Smith	Jones
Value of Another Widget	4	6
Value of Another Getwid	2	2

We will explore production efficiency in more detail in Chapter 16.

EXCHANGE EFFICIENCY

A second condition necessary for an economy to be economically efficient is exchange efficiency. Exchange efficiency exists when there are no more mutually advantageous trades possible. If there is a possible trade that would benefit one transactor and in no way harm the other, then further improvement is possible and the original situation is not economically efficient. Further trade would increase value people perceive, and thus increase economic efficiency.

Like production efficiency, exchange efficiency requires that an equimarginal condition be met. The relative values (or marginal utilities) must have the same ratios for all people. Table 2 shows a case in which this condition is not met. Smith considers widgets twice as valuable as getwids, while Jones considers them three times as valuable. If Smith gives up one widget to Jones in exchange for two getwids, Smith will neither be harmed or helped. The loss of the widget will be offset by the gain of two getwids. Jones, however, will consider himself better off. The loss of two getwids is of less value to her than the gain of one widget (since she considers widgets three times as valuable as getwids). Hence the initial distribution of goods was not exchange efficient. Trading yielded another distribution which the people considered more valuable. Only when the ratio of values is the same for all individuals will mutually advantageous trade be impossible.

If we replaced getwids and widgets with food and heroin in Table 2, the conclusions of the previous paragraph would have been the same. Many people have objections to other people using heroin. However, economic efficiency assumes all goals are equal, that there are no goals better than others or worse than others. If one rejects this assumption, one may not always consider a move toward economic efficiency desirable. Apparently the public in the United States does not believe that all goals are equal, because the law attempts to prevent exchange of heroin, and hence is an impediment to economic efficiency. (It is, however, possible to argue that the use of heroin by one person can affect others who are in no way involved in the transaction. If this argument has merit, restriction of exchange in heroin may not violate economic efficiency. This possibility of an "externality" will be discussed in Chapter 14.)

There are many possible distributions of goods which are exchange efficient. One distribution is that with Jones having everything and Smith with nothing. (Jones may in fact be uncomfortable with this division and give something to Smith, but this possibility is not shown in the way Table 2 is constructed.) Economic efficiency says nothing about fairness or equity.

EFFICIENCY BETWEEN SUBSTITUTION AND TRANSFORMATION

A final condition which is needed for an economy to be economically efficient is that "trading with nature" must have the same "price" as trading with people. If an economy is on the production-possibilities frontier, then more of one good, good A, can be produced only by shifting resources and thus producing less of something else, good B. The **marginal rate of transformation** refers to how much of good B must be given up to produce more of good A. When society moves along the production-possibilities frontier, it is trading with nature, and the marginal rate of transformation is the price it pays to get more of good A.

The amount of good B that must be given to people so they willingly give up a unit of A must be the same for all people if the economy is exchange efficient. The ratio at which people will trade B for A is the **marginal rate**

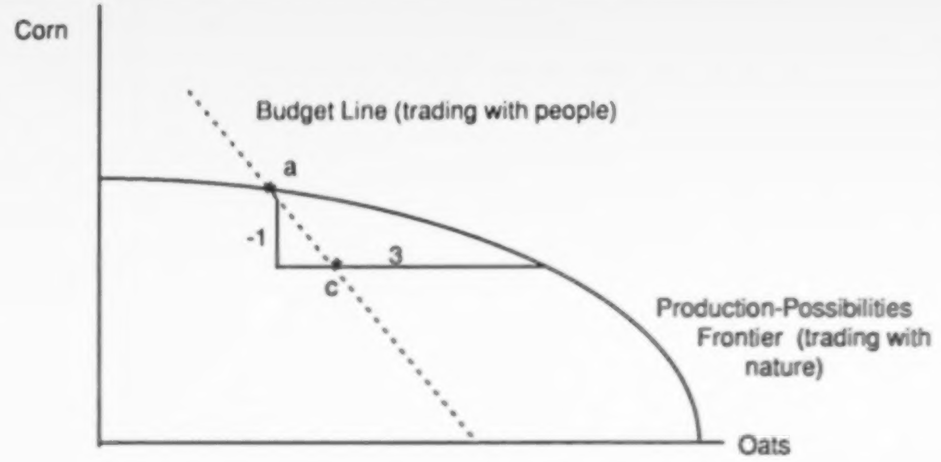

Corn
Budget Line (trading with people)
a
-1
3
c
Production-Possibilities Frontier (trading with nature)
Oats

Figure 11.4 Trading with nature should not be more (or less) profitable than trading with people.

of substitution. Only when the marginal rate of transformation of each product is equal to the marginal rate of substitution, that is, when "trading with nature" has the same "price" as trading with people, will the economy be economically efficient.

A numerical example makes this condition easier to understand. Suppose that producing one more bushel of corn requires the sacrifice of three bushels of oats and vice versa. This statement tells us how the economy can trade with nature. Also suppose that people value corn and oats equally, that is, they will trade a bushel of oats for a bushel of corn and vise versa. Even if the economy is production and exchange efficient, which means that it is on the production-possibilities frontier and that given the amounts of corn and oats produced, all mutually advantageous trades have been made, there are changes which can increase value. If one less bushel of corn is produced, three more bushels of oats can be produced. Since a bushel of corn and a bushel of oats have the same value in the eyes of consumers, the sacrifice of one bushel to get three is a desirable sacrifice. Hence this shift will increase value and thus efficiency.

Figure 4 illustrates the discussion of the previous paragraph. Both point *a* and *b* are on the surface of the production-possibilities frontier. Moving from *a* to *b* means a loss of one unit of corn and a gain of three units of oats. Points *a* and *c* are on the same budget line. Discussion in previous chapters has shown that the slope of the budget line depends only on the relative price of corn and oats, and that all points on a budget line represent equal expenditures of money. Thus point *a* is valued the same as point *c*. Since a move from *a* to *c* leaves consumers with the same value (because they consider a loss of a unit of corn as exactly offset by the gain in the unit of oats), a move from *a* to *b* would increase value. (You should be able to see that one can also argue that point *b* is better than *a* by referring to the marginal costs and marginal benefits of the move.)

A move from *a* to *b* in Figure 4 would change the marginal rate of transformation (the slope of the production-possibilities frontier). It might also change the marginal rate of substitution (the slope of the budget line). As the economy slides down the production-possibilities frontier, the slope of the production-possibilities frontier should get steeper, and the slope of the budget line flatter (because oats should get cheaper relative to corn as people get more oats and less corn). When the two curves have equal slopes, the budget line will be tangent to the production-possibilities frontier. Here the marginal rate of transformation equals the marginal rate of substitution and this point is economically efficient.

The inefficiency of producing too much or too little bread discussed earlier is caused because the marginal rate of transformation does not equal the marginal rate of substitution. We will meet more cases of this form of inefficiency when we discuss monopoly and externalities in later chapters.

In Chapter 7 we saw that economists had reached the paradoxical conclusion that slavery was productive but inefficient. Figure 4 can help us understand this conclusion if we re-label its vertical axis as "leisure" and its horizontal axis as "cotton." Figure 4 now tells us the obvious, that more cotton production requires more work and less leisure. If an employer of free labor wants to get the extra cotton that moving from *a* to *b* gives, he must be willing to pay enough so that the workers will agree to give up this leisure. But with slavery, *force* is an alternative to pay. Preferences of slaves can be ignored.

The key to the productivity of slavery, according to Fogel and Engerman, was that owners were able to extract a large production of cotton, such as that from *a* to *b* in Figure 1, without paying for it. After slavery ended, some planters tried to keep the gang labor which had worked so well with slavery. But they found that when they had to pay workers to work in these conditions, the pay needed to attract workers was greater than the value of production. Only the use of force, which meant that the value slaves placed on good working conditions could be ignored, allowed the superior productivity of slavery. But since the system ignored this value, total value was not at its maximum and the system was inefficient.

QUALIFICATIONS

The discussion of the market and efficiency rests on many assumptions, as we have noted above. Two deserves special notice.

First, the discussion above took place in a *static* framework, where wants, resources, and technology were held constant. But the most distinctive aspect of a market economy is its dynamic nature. Market economies are constantly changing their environments. In contrast, other economic systems, such as tribal-subsistence, feudal, or centrally-planned socialist societies, have a strong tendency not to change. Change is brought in from outside forces, not from the internal workings of the economic system. Karl Marx may have been the first to recognize the dynamic characteristic of capitalism when he wrote in the *Communist Manifesto*:

> "The bourgeoisie has played a most revolutionary role in history. The bourgeoisie, whenever it has got the upper hand, has put to an end all feudal, patriarchal, idyllic relations. ... The bourgeoisie cannot exist without constantly revolutionizing the instruments of production, and with them the whole relations of society. ... Constant revolutionizing of production, uninterrupted disturbances of all social conditions, everlasting uncertainty and agitation distinguish the bourgeois epoch from all earlier ones."

The danger of trying to use conditions of efficiency which were developed in static models for policy guidance in a dynamic world concerned Joseph Schumpeter when he wrote:

> "A system—any system, economic or other—that at *every* point of time fully utilizes its possibilities to the best advantage may yet in the long run be inferior to a system that does so at *no* given point of time, because the latter's failure to do so may be a condition for the level or speed of long-run performance."[1]

---

[1] *Capitalism, Socialism, and Democracy*, 1946, p. 83.

Second, consider Jane and Janet who are very envious of each other. Whatever makes one better off makes the other unhappy. With a system of preferences such as this, the notion of economic efficiency may lose meaning as there may be no possible moves which are mutually advantageous. Economists have generally dealt with this problem by assuming that preferences of this sort are not legitimate and should be ignored.

Unfortunately, people do not ignore preferences of this sort in their behavior. Most social sciences other than economics consider the quest for status a powerful motivator. Status, however, is not like normal goods because it is fixed in amount. If one person gains status, someone else must lose it. Everyone cannot be at the top of the social pecking order. Someone must be at the bottom. Efforts to gain status may help the individual, but must come at someone else's expense. Pursuit of status leads to a prisoner-dilemma situation.

David Hemenway, in his *Prices and Choices* (Ballinger, 1977) tells a fable of a world in which status is acquired by being taller than average (not unlike the world we know) and in which people determine the height of their children shortly after conception. As people strive to improve the status of their children, they ask for taller and taller children. As a result, average height of adults shoots up past ten feet. Eventually in Hemenway's fable the average height has gotten so high that people need braces just to stand up. Clearly the end result is less desirable than the original. Yet this result comes about through a series of moves which individuals take to improve their own well-being. Though Hemenway's story is fiction, there are cases in which quest for status has lead to physical mutilation. An important example was binding of feet in China, which crippled women.

Though the concept of economic efficiency can survive the introduction of status, pursuit of status complicates its use. It adds a large class of situations in which mutually advantageous exchanges may be socially harmful. Economists have not studied these situations in depth. They have found it easier to ignore that the pursuit of status is an important goal for many people.

SUMMARY

Suppose a resource has three potential uses. In use A is produces output which consumers value at $25, in use B it produces output worth $22 to consumers, and in use C is produces output worth $20 to consumers. The concept of economic efficiency says that these resources should be used to produce the output of use A.

Furthermore, there is reason to expect that a market economy will in fact use the resources for use A. Since use A has the highest value to consumers, we expect that the producers of A should be able to bid the most for the resources. Thus the market system should be able to shift resources to their highest-valued use.

The past two paragraphs summarizes the major points of this chapter. In upcoming chapters we will see a number of cases in which the market does not get resources to use A. When the market cannot get resources to use A, economists say that we have a case of **market failure**.

QUESTIONS

1. Suppose that John thinks one banana is worth 4 apples while Sue thinks one banana is worth 2 apples. If they both have apples and bananas, what trade will take place?

2. Suppose that for widgets the MPc is 4 and the MPl is 2. For getwids the MPc is 8 and the MPl is 4. Suppose 1 unit of capital is transferred from widgets to getwids and 2 units of labor are transferred from getwids to widgets. What happens to the output of each?

3. (a)Suppose that for widgets the marginal product of capital (MPc) is 10 and the marginal product of labor (MPl) is 1. For getwids the MPc is also 10, but the MPl is 20. Could the makers of widgets and getwids trade capital and labor and both increase output? How?
(b) Suppose the price of both capital and labor is $10.00, that the marginal revenue of getwids is $.75 and the marginal revenue of widgets is $2.00. Would these prices encourage firms to make an efficient trade of capital and labor? Explain.

4. The final test for Economics 98 is at 3:00 on the last Friday before vacation starts. Barb Bright suggests to Professor Smith that, since it does not matter to him when the test is held, it would be economically efficient for the class to pay him $20.00 and he could re-schedule the test for Wednesday evening. The entire class agrees with Barb that this would be a good idea. (a) Is Ms. Bright right that this move would increase economic efficiency? (b) If she is right that such a transaction would improve economic efficiency, why would this transaction violate the rules of virtually all colleges?

5. (A) The backward economy of Scimonoce has these five farms which produce either corn or oats. Derive the production-possibilities frontier. (Hint: see Chapter 2, Question 9.)

FARM	CORN	or	OATS
A	50		10
B	40		10
C	30		10
D	20		10
E	10		10

(B) Suppose consumers are willing to pay $1.00 a bushel for oats regardless of the amount being sold. (This assumption violates the law of diminishing marginal utility, but makes the problem simpler.) Their demand for corn is given in the willingness-to-pay table below. Does the fact that the 150th unit is worth less than the first unit make sense to you?

The first unit is worth $3.00
The 50th unit is worth $ .75
The 90th uint is worth $ .50
The 120th unit is worth $ .33 1/3
The 140th unit is worth $ .20
The 150th unit is worth $ .05

C) Suppose 50 units of corn are being produced. If 40 more units are produced, what is the minimum value consumers would place on the extra production? (Hint: Use the numbers in the table above.) 40x____=_____.

What is the cost to consumers of 40 more units of corn? (Use the production-possibilities table to find out how much oats will be lost. Then compute the value of these oats in terms of what the consumer would have been willing to pay for them.)

If individual preferences matter, is it better to produce 50 units of corn or 90 units?

D) Suppose that originally 140 units of corn are being produced. If corn production dropped to 120, what is the maximum loss this would cause consumers? What would the value of the gain be from the increased oat production? Is 120 units of corn more in the interests of consumers than 140? (Remember what happens to oats!)

E) If individual preferences matter, what is "socially optimal" (economically efficient) output? (Hint: In part C you compared 50 and 90, and in part D 140 and 120. In the same way compare other points on the production-possibilities frontier in terms of their value to consumers.)

6. In the last question you determined what consumers wanted. In this question you will continue using the same production possibilities and preferences to discover whether producers will give them what they want.

Suppose corn and oats require identical amounts of all resources, that is, there are no costs except the opportunity costs shown in the table above. (Leisure is not a third good that we will worry about.) Further, suppose that farmers can always sell their oats for $1.00 a bushel which provides them with enough income to stay in business. (This assumption simplifies the problem.) Finally, assume that farmers are unaware that their production decisions will have any effect on price.

A) Suppose the price of corn is $.19. In the table below determine how much revenue each plot could make planted in corn and planted in oats. Then decide whether any plot should produce corn.

	Revenue From	
PLOT	CORN	OATS
A	___	___
B	___	___
C	___	___
D	___	___
E	___	___

B) Re-compute the table above if the price of corn is $.21. Will any plot make more revenue planting corn than planting oats? How much oats will be produced when corn sells for $.21?

C) If plot A is used to produce 50 units of corn, 10 units of oats are sacrificed. If 25 units of corn are produced there (by using one half to produce corn and one half to produce oats), how much oats are sacrificed?

If one unit of corn is produced there, how much oats is sacrificed? Given that a unit of oats is worth $1.00, what is the value of oats lost when one unit of corn is produced?

D) How high must the price of corn be before farm B will grow corn instead of oats?

E) How high must the price of corn be before farm C will grow corn instead of oats? Farm D? Farm E?

F) Use the answers you have obtained above to fill in the supply schedule below:

Marginal Cost or Price	Amount of Corn
____	50
____	90
____	120
____	140
____	150

There are two ways to interpret this supply curve. One is that it shows how much corn will be supplied at various prices, that is, it tells us about the behavior of the farmers if they are trying to maximize profits. The second is that the first column tells us the opportunity cost of additional units of corn. This was the interpretation that part C was trying to make you see.

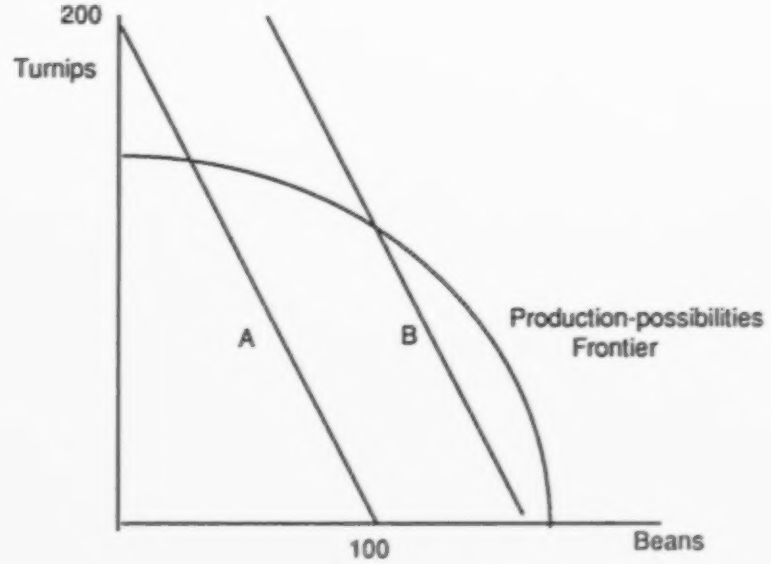
G) Graph the supply and demand curves. (Part F gives you supply, and demand is the willingness to pay curve in question 5.)

H) Suppose the economy was originally producing 50 units of corn. Are there forces which would lead to a change in the amount of corn produced? (Hint: What would the price of corn be? How much corn would farmers like to sell at that price?)

I) Given the supply and demand schedules above, how much corn would a a free market produce? How does this compare to what is "socially optimal"?

7. Lower Watab is a country populated by two tribes, Beaneaters and Turniplovers. Beaneaters, as the name suggests, prefer beans to turnips, and eat turnips only when the price of turnips is very low or when they get very tired of beans, which only rarely happens. Turniplovers, on the other hand, much prefer turnips to beans. As is usual in fictional economies, Lower Watab produces only two products, beans and turnips.

Assume that initially Beaneaters own most of the land (which is the only productive



resource—labor produces nothing). Because they own most of the productive resources, they also get most of the income. Thus the preferences of the Beaneaters will have more weight in determining the value which the society places on beans and turnips than will the preferences of the Turniplovers. Beans will be more highly valued than will turnips.

The graph below shows the production-possibilities frontier for Lower Watab. Budget lines A and B indicate that a ton of beans is worth two tons of turnips. All points on each budget line have the same value given this price ratio. Thus if the economy produces anywhere on budget line A, one can think of the value of production as worth 100 tons of beans. It should be obvious that a position on budget line B is of greater value than any position on A.

a. Draw in the budget line that represents the largest value that the economy can obtain given that a ton of beans is worth two tons of turnips. Can you see how the three conditions of efficiency discussed in the chapter are involved?

b. Suppose that the Turniplovers gain political power in Lower Watab, and begin to tax Beaneaters and subsidize Turniplovers. As a result, the preferences of Beaneaters will become less important and those of the Turniplovers more important. (Sometimes economists talk about "dollar voting," or casting your dollars for the products you want. As a result of the change in income, Beaneaters have fewer "dollar-ballots" to cast and Turniplovers have more.) What should happen to the price of turnips in terms of beans? Draw a budget line on the Figure above to illustrate it. Where will the new position of economic efficiency be? (Moral: economic efficiency will depend on the distribution of ownership of resources. Any change which shifts income from one group to another may also change the efficient allocation of resources.)

8. In this chapter we added up individual demand curves to get market demand curves. In the last section, however, we noted problems can arise when when people have interdependent preferences.

a) John Bandwagon is an insecure individual who bases many of his buying decisions on what he sees his neighbor, J.P. Souza doing. The first demand curve below shows Bandwagon's demand curve for car washes per year if Souza never gets a car wash. The second demand curve shows Bandwagon's demand curve if Souza gets one or more car wash per year. The third demand curve shows Souza's demand curve. Compute what their combined demand curve looks like in the last table.

1		2		3		4	
AMT	PRICE	AMT	PRICE	AMT	PRICE	AMT	PRICE
1	$7.00	1	$9.00	1	$5.00	1	
2	4.00	2	6.00	2	4.00	2	
3	3.00	3	4.50	3	2.00	3	
4	1.00	4	3.50	4	1.00	4	

b) Verily Snob has demand curve 2 for lawn treatments per year if his neighbor does not buy any lawn treatments. If his neighbor does buy at least one, Snob's demand curve shifts and becomes demand curve 1. His neighbor has demand curve 3. What will their joint demand curve look like?

CHAPTER 12

PROBLEMS OF INFORMATION, RISK, AND EXCLUSION

After you read this chapter and work the problems at the end, you should be able to:

Define screening, signalling, speculation, arbitrage, hedging, entrepreneur, moral hazard, free rider.
Explain why a public good will not normally be provided by the market.
Explain two ways of coordinating economic decisions.
Explain how the free-rider problem reflects the prisoners' dilemma.

The past five chapters have developed the core of economic theory. They showed how consumers should act it they want to maximize utility and how businesses should act if they want to maximize profit. Chapter 11 asked if these maximizing efforts on the part of individuals led society as a whole to maximize anything. It found they they did. They maximized value. However, in order to get this result of economic efficiency, we had to make some assumptions.

The assumptions which we needed were the same assumptions we needed to get the supply and demand curves of Chapter 4. All buyers and sellers must be price takers. Information must be good and equally available to buyers and sellers. Property rights must exist and be complete. When these assumptions are not satisfied, the market can fail to allocate efficiently.

This chapter is the first of five to investigate these assumptions and their implications. It consider a series of problems which an exchange economy has because information is expensive or because benefits from an activity flow automatically to others. These problems yield unintended consequences which are important in the way economists view an exchange economy.

WHY ARE THERE BUSINESS ORGANIZATIONS?

A striking aspect of contemporary market economies is the importance of large, organized groups working together as business firms. The individual businessman who is employed by no one and who employs no one is the exception in our economy—only a small fraction of those working fit into this category. There is a puzzle here, because in the large firm the role of the market as a coordinator of decisions is replaced by commands of a hierarchical authority. From a worker's point of view, for example, factory life in the market economy of the United States is very similar to factory life in the command economy of the Soviet Union. In both systems the number of people working at various tasks in a factory as well as the schedule they must work is determined not by a market, but by orders coming from above.

The division of labor provides one possible reason for the existence of organized firms. If a complex task is divided into a series of parts, and a specialist does each part, much more can be produced than if everyone tries to do the whole task alone. Adam Smith wrote a famous passage emphasizing this point:

> "To take an example, therefore, from a very trifling manufacture; but one in which the division of labour has been very often taken notice of, the trade of the pin-maker; a workman not educated to this business (which the division of labour has rendered a distinct trade), nor acquainted with the use of the machinery employed in it (to the invention of which the same division of labour has probably given occasion), could scarce, perhaps, with his utmost industry, make one pin a day, and certainly could not make twenty. But in the way in which this business is now carried on, not only the whole work is a peculiar trade, but it is divided into a number of branches, of which the greater part are likewise peculiar trades. One man draws out the wire, another straights it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on, is a peculiar business, to whiten the pins is another; it is even a trade by itself to put them into the paper; and the important business of making a pin is, in this manner, divided into about eighteen distinct operations, which, in some manufactories, are all performed by distinct hands, though in others the same man will sometimes perform two or three of them. I have seen a small manufactory of this kind where ten men only were employed, and where some of them consequently performed two or three distinct operations. But though they were poor, and therefore but indifferently accommodated with the necessary machinery, they could, when they exerted them-

> selves, make among them about twelve pounds of pins in a day. There are in a pound upwards of four thousand pins of middling size. Those ten persons, therefore, could make among them upwards of forty-eight thousand pins in a day. Each person, therefore, making a tenth part of forty-eight thousand pins, might be considered as making four thousand eight hundred pins in a day. But if they had all wrought separately and independently and without any of them haveing been educated to this peculiar business, they certainly could not each of them have made twenty, perhaps not one pin in a day...."

However, the division of labor need not cause business organizations. It could instead result in a series of different professions. One could imagine the profession of wire drawer, who would sell his output to the wire straighteners, who would sell their output to wire cutters, etc. The market could coordinate their activities instead of a manager. It would raise price for any good that was in short supply and lower price for any good in surplus, thereby encouraging people to produce each step in the proper proportion. In fact in the earliest stages of the industrial revolution, the textile industry was made up of the separate professions of carder, spinner, weaver, etc. and they worked separately in their own homes. Early in the industrial revolution these trades were brought together under the factory roof, and a system of hired workers working under supervision replaced the individual craftsmen.

One could also imagine a system where one person owned the pin making equipment and contracted with people each day to operate the equipment. This method is still used today, as it was used in antiquity, in agriculture. With it the usual employer-employee relationship does not exist. The hired worker can agree to do a specific task, and pay for each task can vary from day to day. In a sense this is the relationship when a modern firm hires the services of an outside consultant, lawyer, or accountant. The person hired is not considered a normal employee because the power of the employer to give orders—to direct his efforts to where the employer sees fit—is very limited. The tasks that the outside person does are agreed to in a market transaction.

A second possible reason for business organizations is that man is a social animal and that he would rather work in groups than alone. A problem with this hypothesis is that people do not just work together in groups, but they work under the supervision and orders of a "boss." Few people profess to enjoy obeying orders.

An interesting solution to this problem was suggested in 1937 by Ronald Coase. Coase argued that "the main reason why it is profitable to establish a firm is that there is a cost of using the price mechanism." This cost may come from the difficulty of discovering what the prices are, but a more important source is in the problems of negotiating a separate contract for each exchange transaction. Instead a simple contract, often for a fairly long term, is agreed to. The exact tasks that the employee is to do are nebulous, but the employee agrees to obey the employer's orders to do a wide variety of tasks.

Coase said that the firm will expand whenever the cost of organizing a transaction within the firm is less than carrying out the same transaction by an exchange in the market. Development of the giant corporation can be explained by the improvement in management technique. This theme has been developed by Alfred Chandler, who examines the growth of the managerial corporation in America in a book entitled *The Visible Hand* (Belknap Press, 1977). The title refers back to the "invisible hand" of Adam Smith. Smith discussed the ability of the market to coordinate the actions of many individuals, each specializing in the production of some product. The visible hand refers to the role of managers in coordinating the actions of many individuals, each doing a specialized task.

The development of the assembly line is an example of a discovery which made coordination by command cheaper. Most people understand that an assembly line breaks a complex task into a series of simple tasks and thus captures benefits from the division of labor. But it also makes monitoring workers easier. The manager can better see how well each worker performs. Since the production process flows from one task to the next, a worker who does a poor job, or a slow job, will quickly become obvious. Poor performance at one task will cause problems further down the line. Further, there is a whole group of workers, the foremen, on the floor who monitor, try to solve any problems which do occur, and make decisions about where particular people should be used.

Coordination in the firm is by command. (Of course persuasion and various sorts of exchanges may also be important.) A manager determines the exact task a person performs. The manager's challenge is to decide how best to use the available resources. Coordination in the market is based on voluntary transactions. The ability of command to coordinate is obvious, while the ability of the market to coordinate is not so obvious. This difference is apparent whenever the economy experiences a problem; there is always a demand that the government "do something." In the 1930s it responded with massive public works. In the early 1970s when the public realized that inflation was a problem, the clamor for action led to wage and price controls. The rapid rise in oil prices in the late 1970s led to a mass of regulations. In each of these cases, and in a huge number of others that could be cited, restrictions are placed on the market. Instead of relying solely on coordination through the market, the government decides to rely at least partly on command. What is remarkable, however, is that in few of the cases of government intervention is there convincing evidence that the government action had its intended effect, and in a great many there is some evidence, though often not conclusive, that the government action had harmful side effects. The difficulty of achieving results through government action may be largely due to the costs of using command—the visible hand—as a means of coordination.

If there were no costs to using command, one would expect the vision (or nightmare) of Karl Marx to be reality today—monopoly capital. In Marx's vision one capitalist would gobble up another until in the end only a few giant capitalists would own and run everything. Then the revolution would come. But this vision of the giant capitalists ignores the costs of giving incentives to workers, monitoring their behavior, and deciding how best to use their skills. It is ironic that the closest approximations to the monopoly capitalists which Marx predicted are the centrally-planned states of Eastern Europe and the Soviet Union. In their attempt to eliminate the market, they have organized their economies as a giant, state-run enterprise. The poor performances in satisfying consumer wants which these economies have experienced is in part due to their use of command in places where the market is a much cheaper way to organize.

Industries which have the ability to easily monitor quality and speed will find coordination by command is cheap. Such industries will generally have increasing returns to scale and thus large firms. Industries in which it is not easy to monitor quality and speed will tend to have many small firms. Agriculture is an example of this second sort of industry. Attempts to farm on a large scale have not been notably successful in the United States. In the Soviet Union and in other socialist countries the structuring of agriculture in large, organized farms has been disastrous in terms of food production. Agriculture cannot be made into an assembly-line process, and the many factors which determine the success of the crop make it difficult to determine responsibility if someone does shirk. As a result, in the United States very large land owners more often rent their land to independent farmers than hire farmers to farm their land.

In the abstract theory of exchange economics, the internal world of the firm does not exist. All coordination is through the price mechanism. However, the existence of large organizations, established on a voluntary basis, points out that coordination through the "invisible hand" requires resources, and that those who find ways to coordinate more cheaply can profit from their discoveries. In a way the existence of large firms is an irony in a market economy. They indicate that a way to succeed in a market economy is to find a better way to coordinate people than market coordination, to replace the market by command.

Perhaps a reason that team sports are so popular in a country which so honors the individual is that they reflect a paradox of the modern world: those groups which cooperate best internally are best able to compete externally. Team sports encourage both cooperation and competition; the two are not incompatible. The often-heard criticism of capitalism, that it encourages competition and thus discourages cooperation ignores the very existence of the large, coordinated enterprises that are most visible in the economy.

SCREENING AND SIGNALLING

When a person buys a used car from a stranger, the buyer knows that the purchase involves risks. The seller may tell him that the car has always worked well, but the buyer cannot be sure that the seller is truthful. It is in the interests of the seller to say these things whether or not they are true. The seller of a "lemon" could say the same thing as the seller of a quality item. Sellers of quality items need a way of communicating information to buyers so that only truthful information will be transmitted.

How do sellers convince buyers of the attractive qualities of their products? Or how do buyers seek to filter out erroneous information? Sellers of high-quality items must find a way to **signal** information about their products, and these signals must be difficult for those selling low-quality products to duplicate. Buyers must **screen** out erroneous information, but allow in truthful information. These problems do not exist in markets where products are simple and easily evaluated. There is little need for this behavior in many agricultural markets, for instance.

One way a seller can signal the quality of its product is by offering guarantees. If a firm offers a guarantee on a poor product, it will suffer a loss. Therefore it is in the firm's interests only to offer a guarantee on a quality product. The guarantee tells potential buyers that the firm believes that it has a good-quality product, and will stake money on that claim.

Another way a firm can signal quality is by building a brand name. A brand name is valuable only if consumers associate it with quality, and the firm can build this association only with time and resources. Once a brand name is established, it is in the interests of the firm to protect it by not offering a poor-quality product with its brand name. When a firm with established brand names does offer a poor-quality product, it usually puts a different name on the products so as not to endanger the public's perception of its brand name.

Signalling plays an important role in the labor market. An employer has little information about a prospective employee, and can hardly expect to get a realistic answer if he asks the applicant if he is intelligent, has leadership qualities, and is responsible. Rather the applicant must try to prove that he has these qualities. A college education is a way of signalling intelligence and perseverance. Leadership can be signalled by extracurricular activities. In fact some students seek leadership positions primarily for their value as ways to signal leadership to future employers The purpose of a resume is to list those activities which will signal attractive qualities to potential employers.

The fact that a college education signals qualities to employers has raised some interesting questions about why people get college educations. A popular answer among economists has been that it is a way of investing in people, or of building "human capital." This means that education increases the productivity of people. More recently some economists have suggested that this view is wrong or at best only partly true, and that college education serves as a way of signalling to future employers. If education is merely a way of signalling, if it is only a complex gauntlet which eliminates those who are not intelligent and do not have perseverance, then the social usefulness of college education may not be very great. From the viewpoint of the student, it does not matter—the benefits are the same either way. Though most economists believe that education both builds human capital and acts as a signal, the relative importance of these two functions is not yet understood.

While sellers signal, buyers screen. Screening is trying to sort bad information from good information. Consider a bank making loans. It tries to make loans only to people who will pay back and to avoid loans to those who will default. It must adopt some rules that separate those who are good risks from those who are bad risks. When one applies for a loan, the bank asks a series of questions about occupation, earnings, and references that it believes will give it the information that it needs.

Screening also occurs in employment. The existence of "old boy" networks is the result of a screening process. If a person wants to hire someone, he will ask those he trusts (the "old boys") for recommendations. Since recommending someone who is unqualified will lower his prestige in the eyes of the other "old boys," there is an incentive for a person to only recommend qualified applicants. (This system of course makes it very difficult for qualified applicants outside the system to get in.) Also, part of the enthusiasm that employers have for graduates of prestigious MBA programs is that the schools are selective in whom they let in. They try to select only those students who have the right combinations of intelligence and personality traits to ensure success in the business world. Thus prestigious MBA programs act as a screening agency for business, and this, as much as what they teach their students, may account for the high salaries their graduates command.

DEALING WITH RISK AND UNCERTAINTY

The future always brings surprises. Sometimes the surprises are nice, but often they are unpleasant. Many people want ways to protect themselves from the unpleasant surprises. They are willing to pay for protection against risk and uncertainty.

Some economists, following the tradition of Frank Knight, make a distinction between risk and uncertainty. Risk exists when a probability based on past experience can be attached to an event, while uncertainty exists when there is no objective way to place a probability on an event. Thus based on mortality in the past, one can say that the chance that a person who is 18 years old will die during a year is about 1 in 500. Predicting the future course of prices, on the other hand, cannot be based solely on the past, and the probabilities that people will assign to various events are more subjective and intuitive. It is difficult in practice to draw a line separating risk and uncertainty, and this chapter ignores the distinction.

Where some people consider risk a problem, others see it as an opportunity. A **speculator** is one who takes risks in the hope of making a profit. Generally a speculator refers to one who tries to forecast future prices, and then bets his money that he is correct. If a speculator expects the price of gold to be higher in a year than it is now, he can buy gold and wait. If he is right, he will make a profit on his action, while if he is wrong, he will lose.

The speculator is widely regarded as someone who contributes nothing positive to the economy. He does not produce anything, but by buying when prices are low and selling when they are high (if he is successful), he smooths the fluctuations in price over time. His purchases at the time prices are low increases prices at this time, and his sales when prices are high helps keep prices from going even higher. The successful speculator performs the useful task of transferring goods from low-valued uses to high-valued ones.

(However, there may be the possibility of destabilizing speculation if everyone bases his offers on what he expects others to offer. In this case the model of contingent behavior from Chapter 2 might apply. This sort of speculation is unlikely in markets in which annual production is large relative to inventories. It is more likely in markets in which there is little or no production, such as in rare stamps, old master paintings, and to some extent gold.)

The development of **futures markets** allows anyone who wants to be a speculator to become one. In a futures market agreements to buy and sell at a future date are made, with the price set when the agreement is made. For example, there are futures markets for most major agricultural commodities. Using this market, a farmer can fix the price at which he will sell his harvest even before he plants the crop, and a miller or owner of a cattle feedlot can fix the price he will pay for grain in the coming year. In fixing these prices with a futures contract, the farmer and buyer of grain reduce the risk they take. (This is called **hedging**.) They are able to do this because speculators are willing to take risk. Without speculators a futures market could not function properly. The benefits which speculators provide others are no part of their intentions. Here is an example of the unintended consequences that economists delight in.

A person involved in speculation is not engaged in **arbitrage**, he is not a **middleman**, nor is he an **entrepreneur**. Arbitrage is buying in a market where prices are low and simultaneously selling in a market in which they are high. There is no risk involved in pure arbitrage. Arbitrage tends to equalize prices in various markets. A middleman is part of a distribution or marketing network. Though frequently disparaged, the fact that sellers are willing to use middlemen indicates that they do perform a useful service. Middlemen generally try to keep risk to a minimum.

The entrepreneur does deal in risk, but while the speculator reduced the risk of those who do not want to bear it, the entrepreneur's risk is of his own making. The entrepreneur is the creative element in a market economy. His presence makes the system dynamic and ever-changing. The abstract theory of the exchange economy is a static theory, one that has an equilibrium. In real-world market economies, change is always present. New products are created, new managerial techniques developed, new ways of producing products and new resources found. The entrepreneur, the innovator, is the source of change. His role can be understood if one looks at Darwin's view of the biological world, in which a species that found a previously unoccupied ecological niche, or that could better exploit one that was already occupied, would prosper. The entrepreneur is searching for unoccupied economic niches, opportunities to make a profit. The search is risky, and often (probably most of the time) the search ends in failure. But when it is successful, it can change the lives of all of us (just as when a new species evolves in nature, it can change the lives of all previously existing organisms). Most of the large corporations are the results of entrepreneurial effort, though they may no longer be performing much of the entrepreneurial function. Schumpeter, who more than anyone before or since, stressed the importance of the entrepreneur, suggested that there were no permanent triumphs in the search of entrepreneurs. In a process which he called "creative destruction," he suggested that all economic niches would eventually be eliminated by further discoveries of other entrepreneurs.

We have discussed how risk can be shifted to another person who is willing to bear it, the speculator. Another way of dealing with risk is by pooling it using insurance. The idea is simple. If one out of 1000 homes will burn each year, and if each person contributes to a general fund 1/1000 of the value of his home, then the fund will have enough (ignoring administrative expenses and the question of whether expensive homes are more or less likely to burn than cheap homes) to reimburse those whose homes burn down.

The size of the insurance industry indicates that people are eager to pay to avoid risk. They pay and get nothing if fortune smiles on them, while if misfortune strikes, they break even because the insurance should just pay back the value lost in the misfortune.

Sometimes, however, people do better than break even when misfortune strikes, and this possibility has greatly interested economists. If, for example, the misfortune costs a person $1000, but insurance will pay $2000, the insured person has no incentive to avoid the misfortune and may act to bring it on. This tendency of insurance to change behavior is called the problem of **moral hazard**. Sometimes moral hazard is dramatic. Fire insurance encourages arson, automobile insurance encourages accidents, and disability insurance encourages dismemberment.

In a story in its December 23, 1974 issue, *The Wall Street Journal* reported a case that illustrates the problem of moral hazard in a striking way:

> "[T]here is the macabre case of "Nub City," a small Florida town that insurance investigators decline to identify by its real name because of continuing disputes over claims. Over 50 people in the town have suffered 'accidents' involving the loss of various organs and appendages, and claims of up to $300,000 have been paid out by insurers. Their investigators are positive the maimings are self inflicted; many witnesses to the 'accidents' are prior claimants or relatives of the victims, and one investigator notes that 'somehow they always shoot off parts they seem to need least.'"

The problem of moral hazard also affects government programs that insure people against misfortune. A variety of programs help people who suffer the misfortune of poverty. Aid to dependent children helps people who suffer the misfortune of having children to raise that they cannot financially support. Unemployment compensation pays people who suffer the misfortune of losing their jobs. Food stamps and public housing help the poor. Yet all these programs also suffer from problems of moral hazard. They increase children borne out of wedlock, unemployment, and poverty.

Moral hazard is the result of maximizing behavior. A person weighs the costs and benefits of an action, and when benefits exceed costs, he takes the action. This does not mean that if a person has a building insured for $50,000 but which only has a market value of $30,000, the owner will necessarily commit arson. There may be costs of violating one's moral code and of getting caught and convicted for arson. But some people put into this situation will find a way to torch the building because they do not find the cost of violating a moral code very high and they consider the chances of being caught small. And other people will be less careful about avoiding fires. Moral hazard does not require that people *intentionally* cause the misfortune. If they simply take fewer measures to prevent misfortune, the same outcome occurs.

The problem of moral hazard creates problems both for private insurance and the government. Private insurance tries to keep the insured value of any misfortune less than the value to the insured person. It tries to keep buildings and autos insured for less than their true worth. In addition, it is usually against the law to create the misfortune that you are insured against. Finally, if the problem of moral hazard is too great, there will be no insurance coverage for the misfortune.

The government can and sometimes does take a similar approach. It can give so little aid to those in distress that it provides little encouragement for people to put themselves in the situation. Of course the government then provides little help for those in distress. As it expands a program to provide more aid to those in distress, it also encourages people to put themselves in distress. If people are paid to be poor, some will become poor. If people are paid to have children out of wedlock, some will. If people are paid to be unemployed, more will be unemployed. Thus government programs that act to insure citizens against some misfortunes have a basic tradeoff that cannot be escaped. Greater efforts to help those in need also increase actions that are considered socially undesirable. Unintended consequences abound in the area of moral hazard.

The insurance industry can also face problems of signalling and screening. People who buy insurance often have a better idea of the risks they face than do the sellers of insurance. People who know that they face large risks are more likely to buy insurance than people who face small risks. Insurance companies try to minimize the problem that only the people with big risks will buy their product, which is sometimes called the problem of **adverse selection**, by trying to measure risk and to adjust prices they charge for this risk. Thus life insurance companies require medical examinations and will refuse policies to people who have terminal illness, and automobile insurance campanies charge much more to people with a conviction for drunk driving.

QUALITY AND PRICE

In the late 1960s a large department store in Madison, Wisconsin had a special sale on electric knives. It priced them very low, but surprisingly, few knives sold. The store then raised the price a little, and the knives quickly sold out. What could explain this strange phenomenon, which clearly is not what a demand curve says should happen?

The most likely explanation for this event is that people often make judgements about the quality of items based on their prices. Usually high quality items are expensive and low quality items are cheap. Thus when people see an item which appears unusually cheap, they suspect that the item has low quality. In the story above, people must have thought that something must be wrong with the knives because the price was so low. When the store raised the price a bit, people no longer thought that the deal was too good to be true.

Quality often depends on price. For example, consider how parents would act if they had won a free vacation and had to hire someone to watch their baby. Would they seek out the cheapest care available? If they love their baby, they should be reluctant to act this way. Those who are willing to work for very little indicate that they are unable to get higher wages. They will probably provide low quality child care. This is a case of adverse selection.

Suppose, however, that our couple has found someone whom they trust to watch their baby, and are negotiating the price, Should they try to negotiate for the lowest price they can get? Maybe not. Maybe the effort the babysitter will take will depend on how much she is paid. People have perceptions about what is fair, and often work less if paid less and work harder if paid more.

If all parents find that lowering price decreases the quality of services they get, then even if there is a surplus of people wanting work as babysitters, price may not decline. Such a market would not act as supply and demand suggests, but would have an equilibrium price which is not market-clearing.

Joseph Stiglitz is the economist who has studied this sort of market most extensively. He has called the equilibrium price which is not market clearing an **efficiency price**. The price is efficient not in the sense which Chapter 11 defined economic efficiency, but in the sense that it is the best price from the viewpoint of individual buyers or sellers in the market.

Efficient prices seem to exist in labor markets and in markets for loanable funds. The babysitting case mentioned above has two reasons that an employer might prefer to pay a price higher than the market demands. A third reason can exist if higher wages cut turnover. If time is needed to train employees, then high rates of turnover cut productivity. Hence high wages may result in lower costs for the employer than low wages.

Banks are major sellers in the market for loanable funds. Suppose there is a decrease in the supply of funds which they have to lend. A supply and demand model suggests that price, which in this market is the interest rate, will rise and force out some of those who want to borrow money. Banks, however, have to wonder what sort of customer is left at high interest rates. Perhaps there is a problem of adverse selection, that higher interest rates leave only those borrowers who are willing to take big risks. If higher interest rates increase the riskiness of the loans the bank makes, it may find that the higher default rates reduce profits. In this case, the bank will find that the best course of action will be to limit the rise in its interest rates, and ration funds among borrowers. In this market there will be a *shortage* of funds at the *equilibrium* price. Some people willing to pay the interest rate will be told by the bank that they cannot borrow because the bank does not have funds available.

Efficient prices (and wages and interest rates) are important when one looks at how economic systems adjust to equilibrium. This topic is central in macroeconomics, and as a result, Chapter 13 in volume 2 will look at the topic of efficient prices again.

FREE RIDERS

Suppose an area has a terrible mosquito problem. An enterprising young man decides that it would be fairly cheap to spray the area ponds and swamps to control mosquitos. (In most species of mosquitos, the larval stage is spent living in water, but breathing air.) He also reasons that most of the people would value a reduction in mosquitos a great deal. However, even if his assessment of the situation is correct, an attempt to form a business will run into a serious problem, the problem of the **free rider**.

To see the problem of the free rider, consider the options open to a "customer" of the mosquito reduction service if the customer is asked to pay $25.00. If he pays $25.00, he will get a reduction in mosquitos (if enough other people also pay to make the service financially viable). If he refuses to pay, he will still get the same reduction in mosquitos (again assuming that there are enough other people who pay so the service is financially viable). Once the service is produced, it cannot be withheld from people who do not pay for it. Mosquitos attempt to bite the first target they find, and are not likely to bite only those who do not pay businessmen who are trying to kill them. Given the options, each person is likely to reason that his $25.00 is unimportant relative to the total amount needed to make the enterprise successful, and thus the results will not depend on his contribution. However, if everyone thinks this way, the enterprise cannot come into being. Hence there is a paradox: though people may want the service a great deal, and though it may be cheap to produce, it may not be profitable to produce. Goods and services of this kind are called **public goods**.

Public goods have two characteristics: they are **nonexcludable**, which means that it is not possible to exclude those who do not pay for the good or service, and they are **nonrival**, which means that one person's use of the good or service does not diminish the amount available to others. In addition to mosquito control, the problem extends to police protection and national defense, lighthouses, the development of knowledge, and television.

There are a variety of ways to cope with the free-rider problem. Sometimes there are alternative ways to provide a service that makes exclusion possible: window screens are substitute, though hardly a perfect one, for spraying mosquitos, and cable and pay TV are substitutes for network TV. Network television is a public good, but it has found that putting in something people do not want—the commercials—allows it to prosper as a business even though it "gives away" its services.

Another way that the free-rider problem can be overcome is through social pressure. Community leaders recognize a problem and organize to work to solve it. Those who do not cooperate can be ostracized or in some way their relations with the rest of the community can be adversely affected. If some of those bothered by mosquitos in the example above mutually agree to form an association to work to solve the problem, those who do not voluntarily contribute may find life as an outcast unpleasant. The $25.00 is contributed not to buy mosquito control, but to buy social acceptance. The way a voluntary group can solve the free-rider problem is by changing what the payment buys. Since it is not in the interests of the individual to pay for mosquito control, the group makes him pay for social respectability. In a sense, he contributes for the provision of mosquito control for the same reason he buys underarm deodorant.

An alternative to the social pressures of voluntary groups is the coercion of the government. The government could provide a mosquito control program, paying for it by taxing citizens. Even when voluntary organization is possible, it may be much easier (and require fewer resources to collect the money necessary for the program) to use the power of the government to provide public goods. Further, in those programs in which force is part of the service provided, as in police protection and the court system, or in national defense, governments have usually been unwilling to allow private groups to operate extensively lest they grow to threaten the government itself.

You might notice that the free-rider problem is related to the problem of the commons discussed in Chapter 2 and the prisoners' dilemma of Chapter 3. The same logic is involved in all. In fact the free-rider problem can be seen as an extension of the prisoners' dilemma from a two-person game to a multiperson game. The best solution as far as the group is concerned involves cooperative behavior. However, each individual can be made better off if he and he alone can cheat on the cooperation. Since everyone has this incentive, there is a strong tendency for cooperation to break down unless added incentives (such as force or social ostracism) are introduced into the payoffs.

Note that the free-rider phenomenon is not always a bad thing. For example, it is the free-rider phenomenon that keeps large numbers of sellers from colluding and charging higher prices to consumers. In this case it prevents the formation of subgroups from working against the interests of the whole group.

SUMMARY

At the end of Chapter 11 we considered a resource bundle with three possible uses. Use A had a value of $25, use B a value of $22, and use C a value of $20. A system of making choices which delivered these resources to their highest-valued use, A, is efficient. Chapter 11 suggested that markets could be efficient.

However, in this chapter we have seen a case in which markets were not efficient. If use A is a public good, a producer of A cannot collect most of the $25 of value which he produces. Hence he will be unable to outbid lower valued uses for this resource bundle. The market fails. This chapter also considered problems markets have with poor information, unequal information, and risk, and how people respond to these problems.

At one time economists assumed that any problem a market had could be solved with government intervention. However, the next chapter will show that systems of political choice also face problems.

QUESTIONS

1. How could one test to see if a college education is mostly an investment in human capital or mostly a screening device? What sorts of behavior on the part of students would be consistent with education as investment in human capital, and what sorts are consistent with it as a screening device?

2. Banks and savings and loan associations have often been accused of "redlining," that is, refusing to make any mortgage loans in certain segments of a city. Why might they redline?

3. Assuming that the probability that a person will die in a year is 1/500, what is the minimum price possible for a term life insurance policy of one year for $1000?

4. Stew D. Much is an enterprising student who is always trying to make a buck. He decides to sell grade insurance to his fellow students. He will pay $500 to any student policyholder who gets an F. Since only 2% of all grades are Fs, Stew computes that he can break even if he charges students $10 per class. To make a profit, he decides to charge $15.

a) Is his arithmetic correct in computing the break-even point? (Ignore other problems.)

b) Will he face a free-rider problem? Explain.

c) Will he face a screening problem? Explain.

d) Will he have a problem with moral hazard? Explain.

5. (a) Why is there no private insurance for unemployment?

(b) There is no market that insures against mentally or physically handicapped babies. Yet when people have children, there is a significant risk that they will have a child with such a handicap. How can you explain this gap in insurance coverage?

(c) In Britain virtually all blood used in transfusions is donated. In the United States much is purchased. Typical sellers include college students, slum dwellers, and drug addicts, all groups which can use the extra cash. The United States has had a much harder time keeping its blood supply free of infectious disease than has had Britain. Why?

6. Bob Saylor notices that ships often wreck on a rocky reef off the coast where he lives. Reasoning that shippers would like to reduce the risk of shipwreck, he tells them that he will build a lighthouse if they will pay him.

a) Will he face a free-rider problem? Explain.

b) Will he face a problem of adverse selection? Explain.

c) Will he have a problem with moral hazard? Explain.

7. The standard typewriter keyboard (the QWERTY arrangement of keys) was developed in the 19th century to prevent typists from typing too fast. When typists typed too fast, the keys would jam. With modern, electronic typing, this problem has been eliminated. Why does not the world switch to a more efficient keyboard? (Hint: Look at the costs and benefits for a single individual making the switch. Recall that a typists may not always type on the same machine. Is there a free-rider type of problem here?)

8. Some life insurance will not pay if the cause of death is suicide within two years of buying the policy, and others will never pay for suicide. Why? No life insurance will pay if the cause of death is a world war. Why?

9. Karl Marx defined classes in terms of property relations because then all members of the class would have common interests. He assumed that because all members of the class had common interests, and because they would be aware of these common interests, he could talk about how classes would behave. In their *Communist Manifesto* Marx and Engels also assumed self-interest: "The bourgeoisie ... has left remaining no

other nexus between man and man than naked self-interests, than callous 'cash payment'."

a) Suppose it is in the interests of capitalists as a class to restrict investment to raise the rate of return on capital. Will it be in the individual capitalist's interest to do so?

b) Suppose that it is in the class interests of the proletariat to work together to seize political power. Will it be in the interests of the individual worker to do so?

c) Marx assumed that self-interest translated into class interests because all members of the class had common interests. Can you see a free rider problem which he and Engels did not see? (Comment: This question is adapted from Mancur Olson's *Logic of Collective Action* (Harvard Univ. Press, 1965), pp 102-110.)

10. A Chicago public television, which gets part of its funds through subscription drives, claims that only 10% of those who watch the station contribute. What is more puzzling to an economist, why 90% do not contribute, or why 10% do? Explain.

11. The judgement of quality by price seems to be important in many markets for services. For example, suppose that there is a surplus of doctors or lawyers. The model of supply and demand suggests that this surplus should reduce price. What problems would a doctor or lawyer face who cuts his prices in order to attract more customers?

CHAPTER 13

THE GOVERNMENT AND ECONOMIC EFFICIENCY

After you read this chapter and work the problems at the end, you should be able to:

Explain why dollar voting may not be economically efficient.
Explain why a one-person, one-vote rule may not be economically efficient in determining the amount of a public good.
Define progressive, regressive, and proportional taxes.
Define tax incidence and shifting of taxes.
Explain why profit-seeking is usually socially useful, while rent-seeking is not.
Show using a graph why an excise tax creates a welfare loss.

Much of the development of economic thought has centered on how a market economy functions. In recent years economists have given increased attention to the role of the government in a basically market economy. The government spends funds or gives funds away (transfers funds), it raises those funds by taxation, borrowing, or printing money, and it regulates a considerable amount of activity. Economists have always known that government activity could have unintended consequences. But in recent years as the government has grown larger, the importance of those unintended consequences, and hence economists' interest in them, have also grown.

There are at least two reasons that investigation of government was inevitable for economists. First, Chapter 1 told you that economics is the study of choice. Though many social choices are made in the market, many others are made through a political process. It is natural to wonder how these systems of **public choice** work, and how their results compare to the results of "dollar voting." Second, economics is built on the assumption of self-interest. In examining exchange we assumed that consumers maximized utility and businesses maximized profits. It is natural to ask if the same motive of self-interest drives legislators and bureaucrats. Can we build a theory of government on an assumption of self-interest? If not, why are those in the government special?

This chapter begins by looking at problems a one-person one-vote scheme has. It then asks what assumption of behavior fits those in government, and contrasts a view which has been termed the "public-interest" view of government with the "private-interest" theory of government. It then examines what the private-interest view of government suggests for taxes and also the economic efficiency effects of one tax, the excise tax. It closes by introducing what is probably the logical conclusion of the private-interest view of government, what economists call rent seeking.

MAJORITY RULE AND ECONOMIC EFFICIENCY

Some economists have suggested that one way to view the decisions that are made in the market is to see them as the result of "dollar voting." Consumers are in a way casting votes when they decide to buy a product. If enough of them cast their dollar votes for the product, the product will be profitable to produce. If there are not enough dollar votes cast for a product, it will not be profitable for a business to produce it, and it will not be produced. We saw in Chapter 11 decisions reached with dollar voting which were not economically efficient. Public goods, for example, will not be produced because of the free-rider effect. In this section we will look at the results of majority voting and find that it too can lead to economically inefficient results.

Not all political voting is based on majority voting. In some elections the winners must have a majority, and a run-off election is held if no one gets a majority. But in most elections only a plurality is needed, that is, the winner is the one with the most votes even if he or she gets less than 50%. In most rules of order a vote to close discussion or to change the rules requires a two-thirds vote. In most jury trials in the United States, a unanimous vote is required to convict a person.

Economists, in fact, have often been attracted by the requirement of unanimity. With such a rule, only moves which harmed no one would be allowed. But for such a rule to work, that is, for any but the most trivial motions to gain passage, the gainers would have to be able to buy off losers. Once payments are allowed, however, a serious bargaining problem arises. Any one person can hold up passage of a bill which benefits everyone until the others pay him. Requirements for unanimous passage simply do not work well in large groups, and therefore they are never used. In order to get the benefits or collective action, individuals must agree to accept some outcomes which they do not

TABLE 1
A VOTING PARADOX

	Most preferred		Least Preferred
Smith's Ordering	A	B	C
Jones' Ordering	B	C	A
Doe's Ordering	C	A	B

want and on which they are outvoted.

Majority voting works better, but it too can have serious problems. One problem is that there may be no equilibrium solution. Consider the voting that results from the preferences in Table 1. If these three people vote to see if policy A or policy B should be selected, they will choose A. (Smith and Doe vote for A and Jones for B.) If they vote to see if policy B or C should be selected, they will choose B. (Smith and Jones vote for B and Doe for C.) Since policy A is preferred to B and Policy B is preferred to C, one might expect that Policy A is preferred to C, but this is not the case. If they vote for policy C or A, they will choose C. (Jones and Doe prefer C to A.)

Even if there is a policy that clearly wins, there is no reason to expect that this policy will be economically efficient. This can be shown using Table 2 which contains the marginal valuations that three people have for a public good. Recall that a public good is one that is non-rival, which means that when one person uses it, this use does not reduce the amount available to others. Thus the total extra value that the group obtains from each unit of the good can be obtained by adding the individual marginal valuations. For example, if one unit is produced, the benefits Tom receives are $18, the benefits Dick receives are $19, and the benefits Harry receives are $50, so the total benefit of the first unit is $87.

Suppose the extra cost of the good in Table 2 is $43.50. If this good is provided by dollar voting in the marketplace, only one will be bought. At a cost of $43.50, neither Tom nor Dick will buy any. Tom, for example, finds it a bad decision to spend $43.50 for something which gives him benefits worth $18. The one

TABLE 2
DEMAND FOR A PUBLIC GOOD

Amount		Marginal Value to		Sum of Marginal
of Good	Tom	Dick	Harry	Values
1	$18	$19	$50	$87
2	17	18	10	45
3	16	17	2	35
4	15	16	0	31
5	14	15	0	29
6	13	14	0	27
7	12	13	0	25

unit which is provided will be bought by Harry because he will find it worthwhile to spend $43.50 to get benefits of $50. However, the economically efficient amount is two because the second provides benefits of $45.00, greater than its cost.

The market operating on the basis of dollar voting is inefficient here, but a political process operating on a principle of majority voting may also be inefficient. Suppose that the extra cost is split equally among the three voters (or that each must pay $14.50 for each unit produced). Tom wants four units produced, Dick wants five, and Harry wants only one. Majority voting will result in four being produced. Tom and Dick will outvote Harry below four, and Tom and Harry will outvote Dick above four. This democratic solution is not economically efficient. The extra value to the group that the fourth unit contributes is $31, but its extra cost is $43.50.

In this simple case of majority voting, the median or middle voter has his wants satisfied best. People on the extremes are unlikely to be satisfied. The inefficiency results because this simple voting scheme ignores *intensity of preferences.*

In modern, complex societies few issues are decided by direct voting. There are too many issues to decide in this way, and the amount of knowledge needed to make those decisions is much larger than any ordinary citizen will want to have. (Recall the discussion of the rationally ignorant voter in Chapter 7.) Yet many societies want the preferences of citizens to determine the course of public policy. To make citizen preferences matter, the society can elect representatives who, in the name of the electorate, vote on many issues. To analyze how voting will take place in this system, we need to make an assumption about what motivates the elected representative.

PUBLIC INTEREST AND PRIVATE INTEREST

One assumption of how elected representatives and others in government act is the public-interest view of government. The public-interest view of government sees government as composed of individuals who are motivated by a desire to serve the public by doing what is "right." The government emerges in this view as an instrument which will (or can) improve society. Unintended and unexpected consequences of government actions do not seem to arise in this view of government.

A public-interest view of government is probably most often a normative theory of government rather than a positive theory. It appeals to those who want to discuss how the government ought to act. This view appears in political rhetoric and in much of the political discussion in the press. When one wants a positive theory, one that explains why governments do the things which they do, the public-interest theory becomes nebulous. No proponent of this view has ever written a scholarly summary of the view. To some extent its opponents have invented the public-interest view as a way of making sense of a variety of often-stated positions. For example, without this invented view, it is difficult to explain the thinking of those who reason that the mere finding of a problem in the private sector establishes a case for government intervention. Even the exact meaning of the term "the public interest" is not clear. Economists often assume it means economic efficiency, but other definitions are possible.

The contrasting view is the private-interest or economic theory. In this theory those who are involved in government have the same motivations that those in the private sector have, that is, they are assumed to be motivated by a narrow concept of self interest: wealth, fame, and power. If what is in the public's interest is also in the private interest of the government decision makers, then the public interest will be served. If there is a conflict between the public's interest and the private interest of governmental decision makers, the public's interest will lose. It is the incentive structure which determines behavior, and bad performance by the government indicates, according to the private-interest theory, a bad incentive structure, not a problem with motives.

In democracies, the private-interest view of government assumes that politicians want to be elected (or re-elected). This is the public sector equivalent to the private sector assumption that businessmen try to maximize profit. Further, just as businessmen must respond to customers, and thus often are forced to do those things which the customers want, so politicians must respond to voters. This need to respond to voters can force politicians to enact those programs for which benefits outweigh costs.

However, decisions of elected politicians need not be economically efficient. Consider what happens if the politician must vote for the three policies shown in Table 3. The table shows the percentage of the population which supports and opposes each policy, and these positions can represent the percentages of the population that is helped and harmed by the policies. Initially one would suspect that a politician would vote against all three policies. However, suppose that the people who favor each policy have strong opinions, and will base their votes in the next election solely on how the politician votes on that one issue. Then it may be in the interests of the politician to vote for all three policies.

TABLE 3
THREE POLICIES

POLICY	%FAVOR	%OPPOSE
A	20%	80%
B	25%	75%
C	15%	85%

One cannot say whether or not voting for these minority positions is economically efficient. It may be that the gains to those who favor them are large and the losses to those who oppose are small. Then this voting through elected representatives is a way of taking into account intensity of preferences. On the other hand, recall that voters are rationally ignorant. They may not strongly oppose the policies because they have not taken the time necessary for them to understand that the policies really do harm them seriously.

Thus there is no guarantee that policies made by representatives pursuing their own interests will be in the interests of the society. Rather the theory of public choice leads one to expect that there will be a bias toward those programs which have visible benefits and hidden costs, and against those which have hidden benefits and visible costs. In the economic view, government can fail just as markets can fail. Just as there is no guarantee that the market will make self-interest serve the public interest, there is no guarantee that the institutions of government will force self-interest to serve the public good.

The economic theory of government may seem a bit extreme because it eliminates any role for ideology or decisions based on selfless motives. It is difficult to believe that these factors play no role in any government decisions. The economic theory is an attempt to build a scientific theory, one that makes risky predictions. If one does not restrict the motives that the actors can have, then there are no outcomes that are impossible, and thus the theory cannot be scientific. The economic theory is an exploration of how much of government activity can be explained on the basis of particular incentives given narrow self interest. If in the end (and the exploration continues by a host of researchers) the economic hypothesis can explain much or most of government activity, the way citizens approach the government and the demands they put on it may be considerably affected.

THE POLITICS OF TAXATION

The activities of the government require resources and the government must in some way obtain these resources. Its method of obtaining resources from the private sector may affect economic efficiency and the distribution of income, though the government may not always consider these effects. If government raises funds by taxation, the price one party in a transaction sees can be different from the price the other party sees. The divergence of price can lead to economic inefficiency for reasons explored in the next section. On the other hand, the proper structure of taxation can correct a case of "market failure." The ability of taxes (and subsidies, which are negative taxes) to change the way people act is the key to their ability to affect economic efficiency.

Taxes can also change the way that the market distributes income. The normative biases of the majority of economists can be seen in the terminology they use to discuss the effects of taxes on income. In a **progressive** tax the tax *rate* increases as the level of income increases. In a **regressive** tax the tax *rate* decreases as the level of income increases. In a proportional tax the tax *rate* remains constant as income changes. A key word in each definition is *rate*. Suppose that people earning $10,000 pay and average of $100 in a tax, and that people earning $30,000 pay an average of $200 in this tax. This tax is regressive in this range of income because the tax rate falls from 1% to .67% as income goes from $10,000 to $30,000.

The variety of unintended consequences affecting both economic efficiency and the distribution of income that will occur as a government tries to raise funds have long interested economists. The study of these effects has been a major topic of a subfield of economics called **public finance**. However, the private interest view of government suggests that the sort of concerns which fascinate economists may not play an important role in the considerations of those who write the laws. The discussion below uses the assumption that the law makers are elected politicians.

Taxes are unpopular with the electorate but are necessary to pay for government programs which are popular with the electorate. A successful politician will be one who seems to give the benefits of programs and not the burdens of the taxes (or one who seems to lift the burdens of taxes without substantially cutting the benefits of programs).

Taxes which are hidden from the electorate should be very attractive to the completely self-interested politician, as should taxes which seem to raise large amounts of money while affecting very few. Taxes can be at least partially hidden when they are **shifted**, which means that the *incidence* of the tax is not borne by the person who visibly pays the government. A major example of shifting involves the corporate income tax. An employee of the corporation writes the check to the government, but he does not feel the burden of the tax since he pays with corporate funds. Neither does the corporation because the corporation is merely a contractual relationship, and as such it cannot see, think, use resources, or feel pleasure or pain. When the government takes resources from the private sector, it must reduce the resources available for real flesh-and-blood people to use. It is not clear, despite much effort that economists have spent on the problem, which flesh-and-blood people have resources diverted away from them by the corporate income tax. It might be the legal owners, the holders of common stock. It may be the employees—both blue collar and white collar—and others who supply resources to the corporation. It may be those who buy from the corporation. It may be all owners of capital because if the corporate income tax reduces the return investors can get from investing in corporations, they will reduce investments there—driving up the return—and increase them in the noncorporate sector—driving down returns. In the complex and interdependent system of a market economy, a change in one place can have effects in other places that at first glance seem totally unrelated.

The **value added tax** (or VAT) is very popular in Europe but is not used in the United States. It is a sales tax based on the value added at each stage of production. Value added is computed as final sales minus the cost of raw materials and unfinished goods. The tax in Europe also allows the deduction of expenditures on capital goods (machinery, tools, etc.). In its effects the VAT is similar to a retail sales tax, only instead of being collected at the retail level, the tax is collected at each level of production. The advantage of the VAT from the viewpoint of the self-interested politician is that the tax is difficult to see. It is hidden in the final sales price of the product. A reason that this tax has not been adopted in the United States is that the state governments make heavy use of sales taxes to raise revenue, and they fear that if the federal government adopts the VAT, the federal government may eventually seriously undermine the abilities of the state governments to raise revenue. They prefer that the federal government stay away from sales taxes and leave sales taxes to the states. As a result, state governors and legislators have strongly lobbied against all proposals for any sort of federal sales tax.

Another tax that is partially hidden is the tax for social security. The amount of the tax that is shown on the worker's pay stub is only one half of what the employee has paid to the government in his behalf. The other one half is not reported to the employee, but in the eyes of the employer is just as much a cost of employing a person as are wages that the employee does see. A reason that the employee's pay stub shows any of the tax is that the social security program has always been promoted as a pension fund, and the amounts paid in as "contributions" rather than taxes. If a majority of people considered them as taxes rather than as contributions, there might be more opposition to this tax because it is clearly regressive.

The whole social security program raises interesting problems of incidence, because not only may the tax be shifted, but so may benefits. The granting of money to retired persons lessens the need for children to support their parents in their old age. Instead of John and Jane directly giving money to their parents, the social security system takes the money from John and Jane and gives it to their parents. The large transfer of money from young to old that the social security system seems to entail may be largely offset by reductions in other payments that would take place in the absence of the system.

Deficit financing and inflation are other ways to hide taxes. Variations on these methods have a very ancient history. In the days before paper money, money was usually in the form of coins, generally gold or silver. Sometimes the government would collect as much of the previous issue of coins as possible, melt them down, add a generous portion of copper, and reissue them. The advantage to the ruler is obvious. If he could issue three coins for every two he collected, he could sizably increase his expenditures without seeming to increase the level of taxation. But since he was diverting resources from the private sector to his use, someone had to suffer reduced consumption. The new money would be more plentiful than the old had been in the past, and hence less valuable. If money becomes less valuable, then other things must be more valuable relative to it—that is, prices rise. The debasement of coinage is well illustrated in the Roman empire in the first three hundred years after Christ. The denarius evolved from a silver coin about the size of a dime into a copper coin the size of a nickel, and in the process it lost almost all value.

Modern governments have more options than did their predecessors. Financial markets and banks allow them to borrow without debasing their money. Since resources are shifted to the public sector, someone bears the burden of the action. If the government borrows more, then higher interest rates will "crowd out" someone who will not be able to pay the higher rates. Since the government borrowing is only one of a great many factors that influence the level of interest rates, and since interest rates affect other variables such as people's wealth (recall that the price of bonds varies inversely with the level of interest rates), it is impossible to identify the precise individuals whose spending is curtailed by the government deficit.

There is, however, a temptation by modern governments to finance some spending by printing money and thus causing inflation. Most countries have a progressive income tax—that is, as income goes up, so does the tax rate. Thus if income doubles, taxes will more than double. With this system, inflation will push people into higher tax brackets. Suppose a person starts at an income of $10,000 and a tax rate of 10% or taxes of $1000. If prices and wages all rise by 20%, he will have an income of $12,000. If his tax bill were $1200, he would pay the same real purchasing power to the government. But if the tax rate at $12,000 is 11%, he will pay $1320 to the government. This means that the government can raise taxes by causing inflation, without ever formally acting to raise tax rates. (However, see question 5 at the end of the chapter.)

Inflation can also produce a tax bonus for the government because of the way it affects wealth. If there is no inflation and a person has $10,000 in a savings account earning $300 a year in interest, the $300 is income because real wealth is increasing. But if the rate of price increase is 6%, then the real purchasing power of the $10,000 is decreasing and the $300 interest does not fully compensate for this. When inflation is taken into account, there should be a negative entry of $566 since the purchasing power of the original $10,000 has shrunk to $9434 in the second year. Thus the owner of this savings account has suffered an actual loss of about $266. However, the tax laws can and often do ignore the effects of changing real values, and the owner of the account may find that he pays taxes even though in real terms (that is, in terms of purchasing power) he has no income at all.

Governments, given the right mix of taxing and spending laws, can have an incentive to inflate. But once they do, there is an opportunity for political entrepreneurs to inform the public of the damage that inflation does and promise policies that will end inflation. Thus the self-interest hypothesis of government is consistent with either inflation or stable prices. This inability of the hypothesis to give a clear prediction is a weakness.

THE EFFICIENCY OF AN EXCISE TAX

One of the most important aspects of a tax in the eyes of economists is the effects that the tax has on economic efficiency. The previous sections suggests that this topic is of considerably more interest to economists than to politicians. This section looks at how one can examine the efficiency implications of a simple tax, the **excise** tax.

An excise tax is a sales tax on a specific item. An excise tax can be a *per unit* tax or an *ad valorem* tax. The first is a fixed amount of tax per item, while the second is a percentage of the value of the item. Excise taxes are a relatively minor source of revenue for most governments, but they can be examined with simple diagrams, and they illustrate the problems which all taxes have.

Figure 1 shows the supply and demand diagrams for a competitive market with constant marginal costs. With constant marginal costs, the supply curve will be a horizontal line. Before any tax is imposed, *q1* is the quantity sold and *p1* is the price. A per unit excise tax imposes a wedge between the price that sellers see and the price which buyers see. In Figure 1 the existence of this wedge is indicated by shifting up the supply curve to *Stax*. The line *S* shows how sellers react to the prices they see, and the line *Stax* shows how the sellers react to the prices which the buyers see.

(Alternatively, the demand curve could be moved downward by the amount of the tax. With this interpretation, the prices along the axis would represent the price excluding the

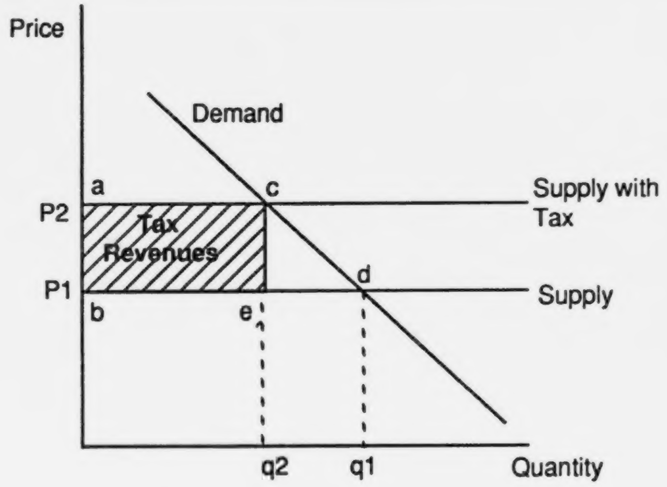


Figure 13.1 An excise tax.

tax. The original demand curve would remain the line showing how buyers react to the price including the tax, and the new demand curve would indicate how buyers would react to the price ignoring the tax. The supply curve would show how sellers react to price excluding the tax, and the intersection of the supply and demand curves would show the equilibrium price excluding tax.)

As a result of the tax, the equilibrium quantity is *q2*. Buyers pay a price of *p2*, but sellers only receive a price of *p1*. The amount of revenue which the government collects is the tax (p2-p1) times q2, or the slashed rectangle.

There are two interesting results shown in this diagram. First, the group which totally bears the tax is the consumers because the price rises by the full amount of the tax. This is true whether buyers or sellers actually write the check which is sent to the government. If the sellers are legally responsible for paying the check, the tax is totally shifted. This extreme result occurs because of the peculiar way in which the supply curve is drawn—it is not a general result of excise taxes.

Second, the tax causes a welfare loss or economic inefficiency because it prevents some exchanges which could benefit both buyers and sellers. There are several ways to show this cost. One is to recall the definition of consumers' surplus. The loss of value to consumers is the loss of consumers' surplus, or in Figure 1, the area a-c-d-b. But the tax revenue is the area a-c-b-e. The difference is the triangle c-d-e. This is a loss to the consumers but it is not a gain to anyone else.

Another way of indicating that there is a welfare loss is to ask what happens to consumers' costs and benefits if another unit beyond q2 in Figure 1 is produced. The extra value of another unit is the distance from the demand curve to the horizontal axis. The extra cost in terms of the value of resources which must be used is the distance from the supply curve S to the horizontal axis. Since this latter distance is less than the former, consumers as a whole would benefit by greater production of the taxed item. But this will not happen because the perceived marginal costs—which include the tax—are greater than the actual marginal costs which includes only the value of resources.

The final way of illustrating the loss of value which an excise tax can cause is with

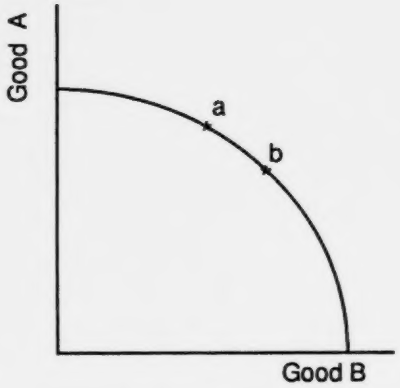


Figure 13.2 An excise tax will, as will a subsidy, shift an economy along its production-possibilities frontier

the production-possibilities curve in Figure 2. Suppose that initially the society is at point *a* and that this point is economically efficient. Now suppose the government places a tax on good A but not on good B. It collects tax revenue, but then returns it to the citizens. The introduction of the tax will reduce the amount of good A people will buy relative to the amount of good B because, with the tax, the price of A has risen relative to B. Hence people will shift from point *a* to point *b* in Figure 2. But point *a* was economically efficient, not point *b*.

Chapter 6 considered the price system as a method of communicating information. The system works efficiently only when the information which is communicated is accurate. Taxes are one source of distorted information. Taxes can distort the signals prices carry so that they no longer correctly describe the costs and benefits that various decisions have for society as a whole. Hence they can distort decisions as well.

Note that the welfare costs of a tax is not the same as the resources which the government uses to collect revenue or the resources which the private sector uses to comply with the tax laws. These are costs in addition to the welfare cost of taxes illustrated above.

The analysis up to this point has avoided complications. In fact, there are at least three complications which one must consider before one tries to draw any conclusions about the actual effect of an excise tax in the real world.

First, Figure 1 is drawn with a horizontal supply curve, implying that resources used to produce the product are all equally productive in producing the item. Often resources are not identical. For example, in producing grain some land is much more productive than other land. If the price of grain drops, some land which is least productive will usually be switched to pasture land or left idle. As the price continues to drop, the land which is taken from production will continue to be the "marginal" land, or that which is least productive (relative to other uses). This pattern of resource productivity will yield an upward-sloping supply curve. (Recall how the differential productivities of land in question 6 of Chapter 11 led to an upward sloping supply curve. It is the opportunity cost that determines the slope, not absolute productivity.)

Increasing costs result in a supply curve such as that in Figure 3. The hatched area above P1 and below the demand curve is the consumers' surplus. The hatched area under P1 and above the supply curve is the producers' surplus, which Chapter 8 discussed.

If an excise tax is imposed in a market such as that of Figure 3, Figure 4 results. In this case the price rises less than the amount of the tax, and tax revenue—the total hatched area—exceeds the loss of consumers' surplus. But there is another loss, one of the producers' surplus. Economic rents are diminished. The total reduction of surpluses will exceed the tax revenue by the amount of the triangular area a-c-b. Hence an upward sloping supply curve does not change the conclusion that the tax causes and inefficiency, though it does alter the conclusion about who bears the tax.

A second complication to our story is that for some cases a simple supply and demand

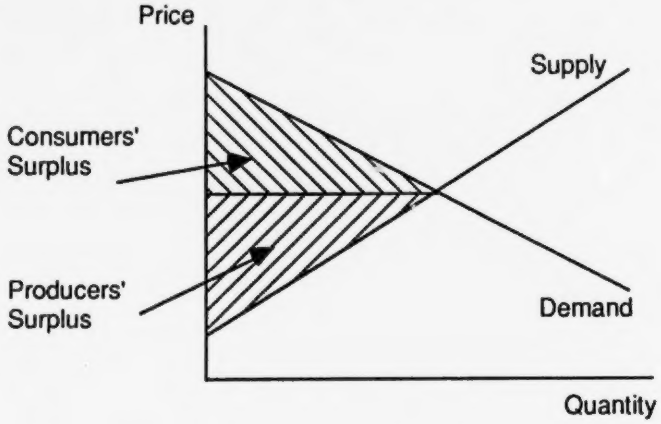


Figure 13.3 The consumers' and producers' surpluses.

graph may not be appropriate to examine the consequences of a tax. Using a supply and demand graph to analyze a tax is a case of **partial equilibrium analysis**. Such analysis assumes that any effects on other markets can be ignored. Sometimes this assumption may not be appropriate. A tax on item A may affect the price of item B, which in turn may alter the supply or demand of item A. When this sort of situation occurs, economists can attempt to analyze it with a much more complicated technique of **general equilibrium analysis**. Though Part 3 uses simplified general equilibrium analysis, its use in

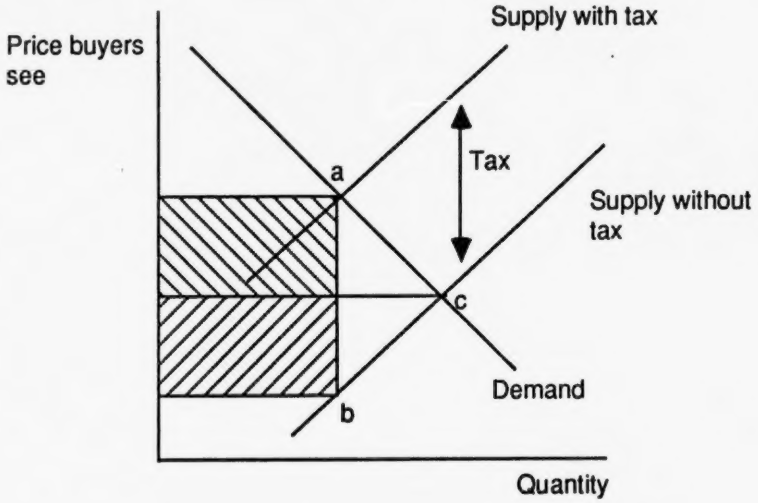


Figure 13.4 An excise tax reduces both consumers' and producers' surplues

examining taxes is best left to advanced courses in economics.

A third complication in discussing the economic efficiency of a tax is that up to now we have been assuming that we started from a position of economic efficiency. The next chapters explore reasons that a market economy may not be economically efficient. If it is not, then the distortion which the tax causes may offset some initial distortion, and, as a result, the tax can increase economic efficiency rather than reduce it.

RENT SEEKING

A **tariff** is an excise tax on imported items. The effects of a tariff are the same as any excise tax: the tariff will raise price, reduce quantity, and decrease the consumers' surplus. A **quota** is a restriction of imports that also raises price, reduces quantity, and decreases the consumers' surplus. For example, in 1985 the world price of sugar was a bit over three cents a pound while the U.S. price was more than six times the world level. This price differential existed because only 2.677 million tons of foreign sugar were allowed into the U.S. A tariff could produce the same price differential, but unlike a tariff, a quota produces no revenue for the government.

The right to sell sugar in the U.S. is valuable. One who has that right can buy sugar at the world price and sell it in the U.S. for six times as much. It would seem that import quotas create a windfall gain for the lucky few who have government permission to import. Sometimes they do. But the situation may be more complex.

Because the right to import—to buy a product cheaply in the foreign market and to sell it dear in the domestic market—is valuable, there should be competition for that right. This competition may take the form of campaign contributions, lobbying, or spending money to prepare proposals submitted to a government bureaucracy. Economic resources are used to obtain this governmentally created right, and as a result, there is not merely a transfer from one group to another as happens with a tariff. The actual social loss not only the triangular area c-d-e in Figure 1, but also the resources spent in competition for the surplus a-c-b-e.

We have here another case of people pursuing their own self interest who contribute nothing to the public good. This competition for a governmentally created right has become known as **rent seeking**. In profit seeking, people are forced to compete in serving the public. Those who best serve the public, by producing the best products at the cheapest prices, receive the most profit. Thus there will be a desirable side-effect to profit seeking. In rent seeking there is no such desirable side-effect.

Rent seeking includes not just competition for the right to import items when there are quotas, but any special governmental favor. For example, various agencies of the government give grants to colleges for equipment to improve teaching. Suppose that a program gives grants averaging $40,000. This money is not given randomly, but is distributed on the basis of a competition of grant proposals. If one in ten proposals wins (a realistic assumption for many grants), and if the average proposal costs $2000 to prepare (also realistic given the value of time of faculty members, professional grant writers, and secretaries needed to prepare a proposal), then colleges will use resources worth $20,000 for each $40,000 they obtain. In most competitions some people or colleges have a much better than a one-in-ten chance, and others have a worse than a one-in-ten chance. If it cost all entrants $2000 to prepare a proposal, the competition described above would not attract anyone with less than one-in-twenty chance of winning.

Rent seeking can occur whenever competition can occur for some existing value. The most interesting cases involve the government because the government can use force in ways that no private citizen or group of citizens can, and that use of force can be made to enrich some at the expense of others. Rent seeking includes competition for subsidies, for special jobs, and for exclusive rights to produce or sell products. Economists often assume that transfers simply redistribute income, with at most a small welfare loss. (See Question 7.) The topic of rent seeking seriously questions this assumption. If rent seeking is important enough in a society—if too much effort of those who have talent and desire is channeled into rent seeking rather than profit seeking—serious harm can be done to a country's growth and development.

Finally, the basic insights of the self-interested view of government and rent seeking suggest that societies based on free markets and democratic governments are unusual and exceptional. Indeed, what is somewhat difficult to explain is how democratic, market societies manage to begin and continue to exist. It would not seem in the interests of those with government power to allow them to begin, but the self-interest view says that it must have been. On the other hand, these insights help explain the popularity of socialism (the absence of significant private ownership of productive resources, which means that all decision-makers are government employees) despite its poor record in improving the lot of the average citizen. Socialism concentrates decisions at the center of a society, and thus increases the power of government officials compared to what it would be with a market economy.

SUMMARY

In their study of how a market economy works, economists concluded that the assumptions necessary to achieve economic efficiency are so extreme that they can never be met in practice. As this conclusion began to emerge in the 1930s and 1940s, many economists suggested that economic efficiency could be attained with the help of government intervention. However, as they studied the ways that the government can work, they began to find that it may be impossible to expect much improvement from government intervention. Utopia is receding and the "dismal science" is sliding back once again to its traditional role as an anti-utopian force. The best of all possible worlds still may not be very good.

The political attitudes of economists can partly be explained by differences in their views of what is desirable and partly on how economists view the unintended consequences that are part of the government's actions. Some see those consequences as pervasive and generally perverse, and believe that since there is little to be done about them, one should try to keep the role of the government to a minimum. Others seem to downplay the pervasiveness of the unintended consequences of government actions, emphasize the undesirable unintended consequences of the market, and stress the good that the direct consequences of government action can have. Still others believe that a purpose of economics is to understand the unintended consequences so that policymakers can better shape government policies to achieve the goals they want. Finally, some economists stress the importance of private interests in shaping government policy and believe economists can have little impact on the workings of government. To these economists policy advice is a waste of time.

QUESTIONS

1. At the end of Chapter 11 we considered a resource bundle with three uses. Use A had a value of $25 to people, use B a value of $22, and use C a value of $20. We said that a system of making choice which delivered the resources to use A was efficient, and argued that markets could be efficient. However, in Chapter 12 we saw a case (public goods) in which markets would allocate inefficiently. a) Let us consider how government policies can affect efficiency. What will happen to the use of the resources in the following cases?

i. The government puts a tax of $4 on use A.

ii. The government subsidizes use C for $6.

iii. The government places a price ceiling of $20 on use A.

b) Suppose public choice replaces market choice. How will resources be used in each of the following cases?

i. Elected representatives make decisions. The beneficiaries of use A are unorganized and inarticulate while beneficiaries of use C are organized and articulate.

ii. Five voters make the decision by majority vote. Benefits from each policy are given in the table below:

		Policies:	
	A	B	C
Citizen 1	5	6	0
Citizen 2	5	6	7
Citizen 3	5	6	0
Citizen 4	5	4	6
Citizen 5	5	0	7

iii. Instead of majority vote, the citizens use rank-order voting. Each of them ranks the options from highest (1) to lowest (3). Then the ranks for each option are added up and the one with the lowest total wins.

2. Three neighbors would like a street light (a public good) installed at their intersection. They have a choice of three models, listed below in the order of cost. The value that each of the neighbors puts on each option is also given. Based on this information, (a) what will the market do, (b) what is economically efficient, and (c) what would a system of majority voting do assuming costs are equally divided.

			Value to	
Option	Cost	Fish	Gill	Pike
A	$180	$100	$50	$70
B	240	100	65	75
C	300	100	60	90

3. A farmer produces $1000 worth of grain. His labor costs are $800, his fuel, seed, and fertilizer cost a total of $100, and he pays $100 for equipment. The miller buys grain for $1000, labor for $6000, pays interest on bonds of $1000 and charges $1000 to depreciation. He produces $10,000 worth of flour and earns a profit of $1000. The baker produces $20,000 worth of bread which he sells to consumers. He earns a profit of $1000, and pays $10,000 for flour, $6000 for labor, and $3000 for fuel. (a) How much would the state collect with a 1% retail sales tax? (b) How much would the state collect from the farmer, miller, and baker with a 1% VAT? Is there any difference between collections with a VAT and a retail sales tax if fuel, seed, fertilizer, and equipment producers also pay a VAT? (c)Prior to using the VAT, most European nations used a turnover tax—sales of each business were taxed even if the sales were to another company. How much would a 1% turnover tax collect? Do you see a way these companies could reduce the tax?

4. Taxes and subsidies are not the only things which can be shifted. An interesting case of shifting occurred with slavery. Draw supply and demand curves for cotton. Now

suppose that slavery is introduced, and that slavery is productive. What will this do to the graph which you have drawn? What will happen to the sizes of the consumers' and producers' surpluses? What do you conclude about the primary beneficiaries of slavery?

5. Only about one in ten submissions to major economics journals is accepted. If it costs $500 to prepare an article for submission, what is the value a researcher places on having his work published?

6. The economics department at St. Germaine College offers a prize of $1000 for the best essay on rent seeking. These students are considering entering:

	Cost of writing the essay	Chances each thinks she has to win
Joan	$100	1 in 3
Mary	200	1 in 6
Karen	150	1 in 5
Kay	200	1 in 2
Betty	100	1 in 20

Which students will enter? Ignoring the costs of judging the competition, how much will be spent to transfer the $1000 to a student?

7. The topic of rent seeking has some interesting things to say about the efficiency of theft. It may seem shocking, but at first glance, there is no economic cost to theft. If, for example, a thief steals $1000 from you, there is no change in value in the system. You, it is true, have lost value, but it is still there, only in someone else's pockets. To make a case against theft on the grounds of economic efficiency, we need to go deeper.

a. A burglar steals goods which you value at $1000, but he can only sell them for $500. Explain why his actions now result in economic inefficiency.

b. To prevent burglars from stealing, you and all your neighbors install locks which cost $100. Explain why such a response makes theft economically inefficient.

c. Suppose the thief treats crime as a business. Instead of doing honest work, he steals. Explain why making crime a business instead of a hobby makes it economically inefficient.

d. If crime is economically inefficient, can a thief find that crime is a profitable occupation?

8. Suppose you are a head of a large government agency. The legislature wants to cut your budget and you do not want it cut. They ask for proposals on how you could cut spending by 10%. What should you do: try to find the least useful things your agency does, or propose cutting the most valuable things your agency does? Explain the rationale for each strategy.

9. Congressman Snaggle helps 500 of his constituents each year get checks from the government to which they are legally entitled, but which have gotten lost in the bureaucracy. These people are very grateful to Congressman Snaggle, and express this gratitude on election day. Dr. Gutnews, a respected efficiency expert, has figured out a way to eliminate the bureaucratic delays which make these 500 seek Congressman Snaggle's help each year. True or False: Congressman Snaggle will be an enthusiastic supporter of this reform measure because it will help his constituents.

10. "Economists often argue that one should not assume that people in government are altruistic. Rather economists argue that these people are as much interested in their own self interest as everyone else. Yet if one looks at high-ranking appointees in any recent U.S. government, one finds large numbers of people who took substantial pay cuts to take the government job. This is inconsistent with the theories of economists and thus proves them wrong." Comment.

11. Suppose that congressmen and senators were selected by lottery (as jurors are) instead of by election. What does the logic of this chapter suggest about their behavior? Would they be more or less honest? Would they be more or less influenced by special-interest groups? In what areas would the role of the government be larger, and in what areas would it be smaller?

12. What effects could the existence of a social security system have on the number of children that a married couple wants to have? Explain in terms of costs and benefits of having children.

13. In 1981 Congress voted to index tax payments beginning in 1985. When taxes are indexed to the rate of inflation, increases in income caused by increases in inflation will not increase the percentage of a person's income that he or she will pay in taxes. Can this vote be explained by the private-interest view? What effect does it have on the benefits the government gets from inflation?

CHAPTER 14

PROBLEMS IN RESOURCE MARKETS

After you read this chapter and work the problems at the end, you should be able to:

Define derived demand.

List three views of what fairness implies about income distribution.

Distinguish between positive and negative externalities and give an example of each.

List three reasons that the market value of a resource can change, and give an example of each.

Explain why economists prefer pollution taxes to direct regulation, but businesses do not.

This chapter examines two important problems which center in the resource or input markets. The first is the question of fairness. The chapter begins by reviewing the forces which determine the amount of income people earn in a market economy. It then looks at several positions about fairness, and follows with a discussion of government policies which change the distribution of income.

The second topic of the chapter concerns what happens when a scarce resource is perceived as being free. This topic expands the discussion of the problem of the commons in Chapter 2, the prisoners' dilemma in Chapter 3, and the free-rider problem in Chapter 12.

MARGINAL PRODUCTIVITY AND INCOME

Chapters 9 and 10 developed the basic theory for resource markets. They introduced the concepts of marginal resource cost (MRC) and marginal revenue product (MRP), and showed that a firm maximized profit when it set MRC equal to MRP. Let us briefly review this result.

Common sense says that a firm will tend to buy a resource if the added benefit to the firm exceeds the added cost. But the added benefit is the MRP. It is found by multiplying the extra output which the resource produces times the value of that output to the firm, which is the marginal revenue. The added cost is the MRC. In the jargon of economics, the common-sense rule above says that a firm should use more of a resource whenever MRP exceeds MRC. Common sense also tells us that if reducing the amount of a resource cuts costs by more than it cuts revenues, a firm can increase its profits by using less of the resource. In economic jargon, the firm should use less of a resource if its MRP is less than its MRC. Then the only point at which it should not change its actions is the point at which MRC=MRP.

The amount of the resource which a firm wants to buy depends in part on the extra revenue it can get from selling production which the resource adds. A resource is valuable not because it directly satisfies some goal, but because it can indirectly satisfy a goal. Economists say that the demand for a resource is a **derived demand**. The demand for the resource is derived, or comes from, the demand for the goods and services which the resource produces.

To push the analysis further, assume that many firms buy the resource and that none of these firms can noticeably influence price in any market. This assumption that firms are price takers means that marginal revenue equals price of output, and that the marginal resource cost equals the price of the resource. To make the following discussion less abstract, suppose that the resource in question is hours of labor, which means that the price in this market will be wages.

With these assumptions, we get a diagram similar to a diagram in Chapter 11. On the left we have Peter Milloy, a representative worker who sells hours of unskilled labor. His supply curve is drawn so that it slopes upward. But he is only one of a great many workers selling in this market. To find the market supply curve, one must add up all the hours of work which will be offered at each wage.

On the right we have Dynamic Enterprises, one of a great many buyers of unskilled labor. Its demand curve will be the downward sloping part of its MRP curve. The MRP curve is the demand curve for labor because the firm hires labor until the point at which MRP=MRC. Hence, given a wage, one can tell how much labor a firm demands by looking at the MRP curve.

Dynamic Enterprises is only one of a great many buyers of labor. To find the market demand, one must add up at each wage the number of labor hours demanded by each firm. The intersection of the market demand and market supply curve gives the equilibrium wage.

There are some important results which

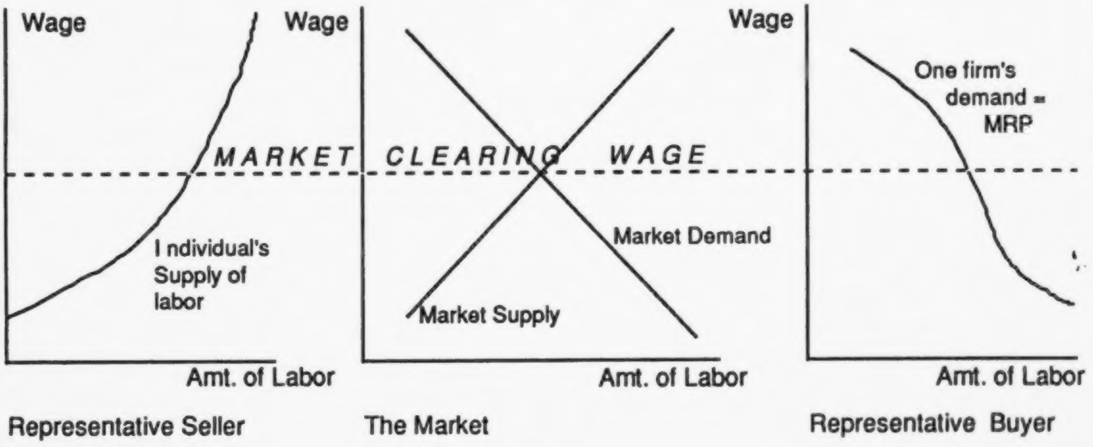


Figure 14.1 The labor market.

come from this simple analysis. First, owners of resources are paid their *marginal* contribution to output. An owner of a resource which *adds* a more valuable contribution to output is paid more than an owner of resource which adds a less valuable contribution. The key words here are "marginal" and "adds". For example, garbage collectors perform a vital service in large cities, and when they go on strike, people are not only tremendously inconvenienced, but their health may be threatened. The total value of the services of garbage collectors is very high. But this high total value does not mean that the value of another garbage collector is high. If adding or removing one garbage collector changes the value of garbage collection services little, then this *marginal* contribution will give a low wage in this market. This analysis exactly parallels the diamond-water paradox discussed in Chapter 8.

Figure 1 can help illustrate changes which will raise or lower wages or other payments to inputs. There are two reasons the demand curve could shift. First, it could shift because people value the output more. For example, the high salaries of professional basketball players is due to the popularity of basketball as a spectator sport. A century ago a man with the same abilities of a Larry Bird or a Michael Jordan would have earned a normal income because there was no demand for his special talents.

Second, the demand curve could shift because of increased productivity, perhaps caused by changes in the amounts of other resources. A trend in industrialized societies has been to replace physical work with machines. As a result, the rewards going to muscle power have probably fallen as compared to the rewards going to brain power. Two hundred years ago a strong but stupid man would probably have earned more than his smart but lame brother. Today the results would probably be reversed.

The supply curve can also change, and as it does, so will the income which the resource earns. A new discovery of metal ore will reduce the value of known deposits. Increased availability of education increases the numbers of doctors, lawyers, teachers, and accountants. If demand for their services does not change, their increased numbers will cause their incomes to drop. People who want to control the wages they earn have usually tried to control the supply curve. The next chapter examines this strategy.

Adding together all buyers and sellers of unskilled labor, or aggregating them, is useful because it gives us insights. Sometimes extreme aggregation is useful, such as treating all labor services in the same market. This sort of extreme aggregation is basic to volume 2. In the case of labor markets, such aggregation is useful if we slightly alter Figure 1.

Equilibrium in the labor market requires that MRC=MRP. With the assumption that everyone is a price taker, this condition can be written as:

(1) Marginal Product x Price = Wage.

If we divide both sides by price, we get:

(2) Marginal Product = Wage/Price = Real Wage.

The real wage, or wages divided by prices, measures the purchasing power of what workers earn. If, for example, wages increase by 10% and so do prices, then the amount that workers can buy has not changed. (Chapter 18 in Part Three explores the notion of real wages more fully.)

Figure 2 graphs the marginal product of labor and the supply of labor as a function of the real wage. Suppose the amount of labor declines. The supply curve will shift to the left, the marginal worker will become more productive, and the real wage will rise. Labor will become more scarce relative to land and equipment, which have not changed. The Black Death in 14th century Europe reduced population by about 30% in a few years. As Figure 2 suggests, real wages rose as a result.

Figure 2 also suggests why economists in the 19th century so often saw population control as a way to solve poverty. Fewer people will, other things held equal, increase wages and more people will reduce them. However in the real world those other factors have often not remained equal. Europe is densely populated and Africa is sparsely populated, yet Europe has a vastly higher standard of living than Africa.

WHAT IS FAIR?

Chapter 6 showed that an economic system must ration the goods and services which are available (and that amounts available might depend on the system of rationing). In a market economy the ability of people to

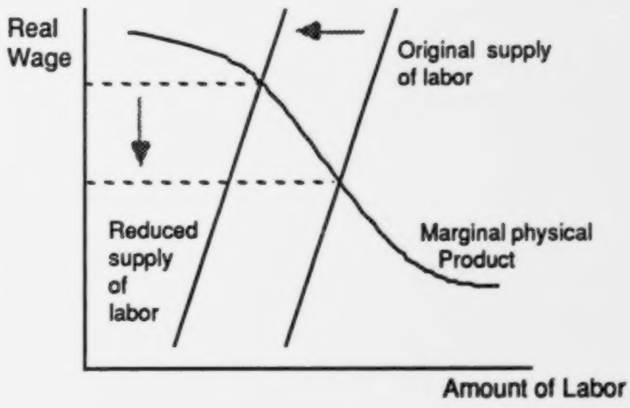


Figure 14.2 The real wage depends on labor's marginal product.

obtain goods and services depends on the marginal productivity of the resources they hold. The most important resource is a person's ability to work, or human capital, but others are ownership of natural resources and capital. Those who hold resources which are highly valued will earn large incomes, while those who hold no valuable resources earn little or no income. This unequal distribution of income which a market system produces raises questions of whether or not a market system is fair.

Fairness is a normative issue, which means that it involves judgements about what is good and what is bad. As a result, economists cannot claim special expertise when discussing this issue. They often rely on arguments from philosophers when they discuss this issue, and they hold widely diverse beliefs. There are, however, some insights from economics which can be useful when one does want to decide issues of fairness.

Unequal incomes do not necessarily mean that the mechanism for producing those incomes is unfair. People have different goals. One result of this difference is that people with identical incomes spend their money in different ways. Some enjoy frequent vacations in Europe, while others prefer to possess expensive cars. Some worry about their homes, while others prefer to spend on entertainment. Another result of different goals is that people with identical abilities may earn different incomes.

Some people have goals which can be met only if they earn high incomes. Other people have goals which require less income but more leisure. Recall that a fundamental trade-off which people face is a leisure-income trade-off. If Mike Fleming wants more income, he must work more, which means he sacrifices leisure. People who value their leisure time highly will earn less income than people who put little value on leisure, other things equal. In a world in which everyone had equal abilities, people would earn unequal incomes if their goals were not identical.

Of course we do not live in a world of equal abilities. Some people are smarter, more athletic, more social, or in some other way more talented than others. Some people, whether through bad genes, bad nutrition, or bad cultural environment, cannot cope by themselves in a complex world. Hence, even if everyone had the same goals, people would earn unequal incomes.

If you look at a situation and decides that it is unfair because one person has too much and another has too little, you probably are making a judgement which compares goals. The judgement says that the person with too much is satisfying goals which are less worthy than those of the person with too little. We commonly make this sort of normative judgement; our decision to give money to one charity rather than another indicates that we find some goals more deserving than others. Our decision to give at all suggests that we decide that the goals of someone else are more worthy than our own "selfish" goals.

Economic analysis suggests that people earn different amounts of income both because they have different goals and different abilities (or "resource endowments," to use a more comprehensive but also more abstract term). From this starting point, we can examine a few common judgements on fairness.

One view is that fairness means everyone should have an equal opportunity to succeed. In this view the process matters, not the results. This position sees economic life as a race. In any race, some people are faster than others. As long as all contestants face the same rules, the race is fair even though some win and others lose. In the economic game some people fail and have low incomes because they made mistakes or were unlucky or did not have enough ability. Yet their failure does not mean that the system is unfair provided that no one erected obstacles in their path. This view is sympathetic to a market system

A second and completely different view is the **egalitarian** position. Egalitarianism judges the results, not the process. It argues that more equality of income is always better than less, and that the best of all possible worlds is one of complete equality. The egalitarian view tends to be unsympathetic to market systems because they generate very unequal incomes.

An influential justification of the egalitarian positions has been developed by John Rawls. Rawls uses the notion of the *social contract*, an idea which Thomas Hobbes and John Locke made famous in the 17th century. Rawls asks what rules we would agree on if we were designing a society which we then had to join, assuming that we had no knowledge of how successful we would be in that society. He argues that most of us would want to join a society which ensures a great deal of income equality.

Both these views have internal inconsistencies. Obtaining equal results requires the use of government power, and only some will be able to wield this power. Those who have the jobs of equalizing incomes will have more power than those who do not; equal incomes results in unequal political power. Obtaining a system of completely equal opportunity is impossible because the results that one generation obtains help determine the starting points of their offspring. People who do well in the economic game will try to help their children succeed by giving them a good childhood environment, a good education, and inherited wealth.

A third viewpoint suggests that income should be determined on the basis of need. This view has a long tradition, going far back in Christian social thought, and has been an important judgement among socialists. A position that equal work deserves equal pay, a position consistent with the opportunity approach, is inconsistent with the need approach.

To implement the need approach there must be some way of measuring need. This measurement is most practical in small group situations, that is, within groups where members know each other well and where members have the same goals. It is hard to implement in large groups of strangers who do not know each other well and who may disagree radically about which goals are worth attaining. One escape from this problem has been to assume that everyone's needs are identical, which collapses this point of view into egalitarianism.

Modern societies have taken aspects of all three viewpoints and established them as public policy. Income taxes are progressive, that is, they take greater percentages of income from those with big incomes than from those with small incomes. This policy can be justified from an equal results point of view. Employment laws require equal pay for equal work. The employer is prohibited from taking factors such as personal need of an employee into account in establishing pay. These laws make sense from an equal opportunity point of view. Finally, tax laws and some transfer payments favor families with more children and higher medical bills. The need viewpoint can justify this aspect of taxes.

Economics cannot tell you what view of fairness is correct. You must make up your own mind about how important various characteristics are. For example, would you rather live in a society with a great deal of inequality but with a lot of social mobility, or in one with less inequality but also less mobility? How important is status? Would you rather have an income of $30,000 and live in a society where this income put you at the top of the social pyramid, or would you prefer another society in which you would be in the middle and earn $45,000? Decisions about fairness involve these trade-offs and several others.

GOVERNMENT REDISTRIBUTION

The views of what is fair are important because a major function of present-day governments involves changing the distribution of income from what the market would give. The discussion in the previous section gave one reason why governments take from some and give to others. People think that the unmet goals of some, the needy, are more important than the met goals of others. People can use the power of the government to change this situation. If people feel strongly about such issues, supporting policies which alter the distribution of income will be politically profitable to elected politicians.

Another reason for government redistribution is that the government can act as an insurer against a variety of misfortunes. Problems of screening (adverse selection) or moral hazard may make market provision of some insurance unfeasible. In addition, once governments begin to redistribute to help the needy, redistribution can have many insurance properties which will increase its political popularity. A program to help the poor, for instance, means that average citizens will not have to worry as much about the possibility of falling into poverty. The government will provide a "safety net."

Governments have emphasized these two reasons for government redistribution, to help the needy and to insure ordinary citizens against misfortune, to justify their extensive redistribution. However, there are two negative aspects to redistribution which are equally important. The first is that transfers change the incentives in a market system, and tend to encourage nonproductive activities and discourage productive ones.

The title of Arthur Okun's book, *Equality and Efficiency: The Big Tradeoff* (The Brookings Institution, 1975) captures the essence of the first problem. Okun, who was close to an egalitarian position, saw that government policies which reduced income inequality could reduce total production. If the government tried too hard to take from the rich to help the poor, the rich would react by changing the way in which they used their resources, and some of new the ways would not be productive. Hence trying to divide the pie more equally could shrink the size of the pie. In addition, economic growth depends in part on decisions which are very risky. If there are only small potential rewards for those who take the risks of innovation, people will stick to the tried and true and economic growth will slow. The problem for government policy in this view is to find the proper balance between income redistribution on one hand and production and economic growth on the other.

Charles Murray has developed a more pessimistic view of the trade-off problem in his *Losing Ground: American Social Policy 1950-1980* (Basic Books, 1984). Murray relies on the idea of moral hazard to argue that the development of the many transfer payments of the welfare state in the United States has *increased* the amount of poverty. If poverty becomes attractive because the government pays well for it, it becomes more common. Murray's contention that there is little room to trade-off equity and efficiency is controversial. Continued study of the extensive programs of government redistribution that most industrialized nations have adopted should, in the next decades, show him right or wrong.

The second negative aspect of government transfers is that people can use government as a way to enrich themselves at someone else's expense. The economic theory of government explained in Chapter 13 emphasizes this third reason for government redistribution.

The economic theory has some interesting predictions which can be checked against the way the world works. One prediction is that the poor will not be especially successful in obtaining money from government transfers. In a democracy transfers will tend to go to those who are organized, who vote, and who are articulate. The poor are none of these.

When government becomes involved in redistribution, organized interests will be at an advantage in obtaining the fruits of that redistribution. They will be able to develop a reason why they need funds, present that case to the government and the public, and keep their members informed about how the government treats their proposals. Individuals who share an interest may not be able to pull together and organize because they can face a free-rider problem. For each individual the costs of joining a group are noticeable, while his effect on the success or failure of the group is not because each is such a small part of the entire group. These unorganized people who share an interest will be less effective in persuading elected officials to aid them.

This organization problem can be overcome if someone finds a good or service to offer that makes membership attractive. For some occupational and business groups, a key for organization is information which will be useful for increasing profits. Other groups form around special marketing services or insurance. Once these groups get established, they often find that their members will benefit if they also engage in political action. The United States has hundreds of special interest groups which lobby the congress for tax breaks and programs to funnel government spending and grants to them. Except for groups with small memberships which are sometimes easy to organize, few of these special interest groups began with lobbying as their major purpose.

Government tends to give money to those who vote and who are articulate. For example, consider the large amount of support which governments give colleges and universities in the United States. State governments often fund more than three-fourths of the cost of tuition at state schools, and the federal government also spends substantial amounts in support of higher education. Only about one fourth of young people receive a college degree, and they will form the bulk of those who in the future will earn high incomes. Further, most of these young people come from families which presently have incomes that are above average. In rich suburbs virtually all students finish high school and often about 90% continue on to college. In poor inner city areas sometimes only half of the students finish high school and only a few percent go on to college. College education represents a case in which those who benefit from state subsidies are, on the average, richer than the taxpayer who funds that activity. State support of higher education takes from the poor and gives to the rich.

Proposals to cut state support of higher education are politically unpopular and not often enacted. Parents of college students, and parents who expect their children to be college students, are much more likely to vote than parents of children who will never go to college. College teachers make up one of the most articulate groups in the country, and can present persuasive reasons why government should increase support for higher education and make cuts somewhere else.

(Since the reader of the above paragraphs is probably a college student who believes that the government should support his efforts to get a college education, a couple of reminders are in order. First, people who receive subsidies are almost never able to take an objective view of them. Second, it is possible to get too much of a good thing. One can look at subsidized education in the same framework as subsidized bread, examined in Chapter 11. Third, the argument that tuition should be kept low so that poor students can afford to go to college does not stand up under scrutiny. A more efficient way to help the poor (who are taxpayers as well as college students) would be to charge high tuitions but have a generous scholarship program based on need. Finally, one can argue that there are "external benefits" of the sort which the next section discusses which can justify large subsidies to education. This argument maintains that a society with a large number of college-educated people is a better place to live in, even for those who do not go to college, than one with only a small number of college-educated people. Some people believe this argument and others do not. However, this argument is built on economic efficiency, not on fairness.)

Finally, the private interest view of government suggests that societies with considerable social mobility will be rare. In a society with a lot of social mobility, poor and unknown people can rise up the social ladder and become rich and famous, while the rich and famous can fall down the ladder and become poor and forgotten. Social mobility looks good for those on the bottom, and it is desirable from the equal opportunity view of income distribution discussed in the previous section.

Social mobility does not look so attractive for those who are at the top of the social ladder. A rich and famous person wants to stay that way, and he wants his children to be rich and famous too. This desire creates an incentive to try to prevent social mobility. Though this desire most easily results in action in nondemocratic societies, it may also limit what one can expect from redistribution policies in democratic societies.

Aristocratic societies, which divide people into commoners and aristocrats, have been historically common and severely limit social mobility. The caste system of India and the apartheid system of South Africa are other structures which have developed to protect those at the top. The Soviet Union, which has equality as its official doctrine, had gradually formed class divisions, and these divisions seem to be getting more rigid. Those who hold power, the *nomenklatura*, are identifiable by the privileges which they have, such as the opportunity to shop in special stores which are off-limits to ordinary citizens. The desire of the *nomenklatura* for their children to stay at the top has made moving up increasingly difficult for ordinary citizens.

It is hard to make a case that a market system distributes income so that only the most worthy goals and needs are met. The market distributes income on the basis of people's contributions to production, and people have widely varying abilities to contribute. There is little reason to believe that ability is related to need. On the other hand, there is not much reason to believe that the government will redistribute on the basis of need. It is naive to view the government as a benevolent dictator which always works in the interests of its citizens. The incentives which guide government actions are as real and as important as the incentives which guide business and consumer actions.

"FREE" RESOURCES

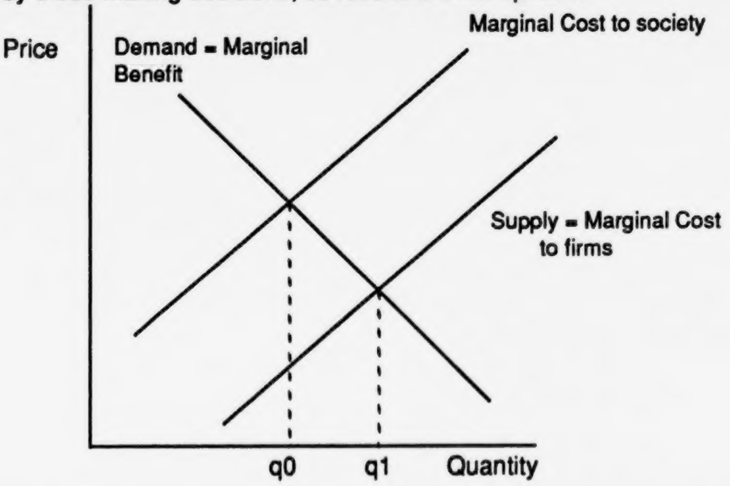
Economists assert that we live in a world of scarcity. We have limited resources, and the cost of using them in one way is that we cannot use them in some other way. There are a few exceptions. Salt water is so abundant in the middle of the ocean that all uses of it can readily be met. Ice in Greenland is also not scarce. Air is our most important free resource. We all get as much as we want to breathe and we do not have to pay for it.

Sometimes resources appear to be free when they are in fact scarce. When scarce resources appear to be free, people do not take into account some effect that their actions have on other people. Because some part of the effect of their actions is "external" to the decisionmaker, these cases are called **externalities**. Externalities can be positive, which means that there are by-products to a decision which benefit others, or they can be negative, which means that there are by-products which harm others. Public goods, discussed in Chapter 12, are an example of positive externality. The problems of the commons and pollution are examples of negative externalities. Because decision involving externalities are based on faulty cal-

TABLE 1
THE FANTASTIC FISHING GROUNDS

NO. OF BOATS	TOTAL CATCH	AVERAGE CATCH	MARGINAL CATCH
5	200 tons	40 tons	—
6	228	38	28 tons
7	252	36	24
8	264	33	12
9	270	30	6

Figure 14.3 With externalities some effects of decisions are ignored by those making decisions, so results are not optimal.



culations of costs and benefits, they generally are economically inefficient.

There are two efficiency problems which arise when scarce resources appear to be free. The first is that less may be produced than could be produced with a different pattern of resources use. The economy will, as a result, produce at a point inside of its production-possibilities frontier. The problem of the commons can have this result.

For example, consider Fantastic Fishing Grounds which are open to all. Suppose that Table 1 shows how "production" of fish depends on the number of fishing boats using Fantastic Fishing Grounds, and that if a boat goes to other areas in which no one else is fishing, it can "produce" 20 tons of fish. John Smith is a boat owner, and he must decide whether or not to go to Fantastic Fishing Grounds. If there are already seven boats there, he will find that it is in his interests to go there too. His choice is between a catch of 20 tons elsewhere or 33 tons at Fantastic Fishing Grounds. But from the point of view of society, this is a poor decision. Adding the eighth boat to Fantastic Fishing Grounds adds only eight tons of fish to the total catch, while sending the boat elsewhere would add 20. The decision to go to Fantastic Fishing Grounds looks good to John Smith because he does not take into account the lost "production" that his decision causes other fishing boats.

The second problem when resources are "free" is that the wrong mix of goods and services will be produced. In terms of Chapter 11, the marginal rate of transformation will not equal the marginal rate of substitution. This is a common result when decision-makers do not take into account some by-product of their actions which burdens or benefits others. A polluter, for example, considers air or water free. For him dumping pollutants into the air or water is a cheap way of disposing of wastes. Yet his actions do involve costs since he affects the alternatives that others face. The polluter may make others forego clean water. One could say that the polluter imposes some costs of production on others, though this use of the word "cost" differs from the normal meaning of cost. Those who bear this cost are not involved in the choice, and in its pure meaning cost is an alternative foregone in a choice.

It is easy to show that when a decisionmaker ignores some costs of his decision, his decision may be economically inefficient. Figure 3 assumes that the market can be represented by supply and demand curves. The demand curve represents the marginal benefit to consumers (and to firms since they are price takers). The supply curve represents the marginal cost to sellers, and because producing the product requires resources that could be used elsewhere, it also represents a cost to buyers. But the production of the product also generates an unwanted by-product which sellers ignore. The marginal cost from the point of view of society as a whole includes this by-product and is thus higher than it seems to the firm. The economically efficient amount to produce in Figure 3 is q0, but the forces of the market will tend to result in the production of q1. (Question 1 at the end of the chapter gives you the same analysis in the form of a table.)

If negative externalities cause too much of a product to be produced, positive externalities should cause too little to be produced. When a person improves his house, his neighbors benefit. Because the decisionmaker will not generally consider these spillover advantages to others, less than the efficient amount of the activity will take place. In terms of a supply and demand diagram, the marginal benefit curve as perceived by the decisionmaker will be to the left of the marginal benefit curve of society as a whole, and thus too little of the activity will take place.

When scarce resources are perceived as "free", there will be potential value that a market will not capture. Is it possible for a society to capture this value, and if so, how?

A common "solution" to this problem has been to assume it away. This solution is especially common in plans for utopias, and writers in the Marxian tradition frequently illustrate it. Pollution exists, in some of these arguments, because capitalistic man is greedy, but when the new socialist man comes into existence, the problem will cease. Solution by assumption has at times crept into mainstream economic thinking as well.

A more practical solution is a system of property rights. This is an ironic solution in a way, because many environmentalists and ecologists have argued that the existence of externalities proves that a market system is seriously flawed and should be scrapped for an alternative, generally with greater state ownership and control. Economic analysis, however, shows that externalities exist when property rights are incomplete. Reducing the role of private property would make the externality problem worse. When no one owns the air or water, there is no incentive to avoid an overuse of the resource.

In the 19th century buffalo were hunted almost to extinction. The cause was a classic example of the problem of the commons. If an individual hunter limited his kills, it is unlikely that he would have benefited from his restraint. The buffalo which he did not kill would probably have been killed by someone else. Yet from the point of view of buffalo hunters as a group, the optimal strategy would have been to limit killing so that the industry could maintain itself indefinitely.

In contrast to the buffalo, the number of cattle in the American West increased during the 19th century. The key difference between the different fates of buffalo and cattle was not that buffalo hunters were greedy and cattle raisers were not. It was that cattle were privately owned, and buffalo were "free." Property rights force people to take into account all costs and benefits of their actions. A cattleman's decision to kill or not kill his cattle did not affect other cattlemen in the way that a buffalo hunter's decision to kill or not kill buffalo affected other buffalo hunters. When a resource is owned by all, when it is "free", there is a strong tendency for individuals to misuse that resource.

The existence of property rights allows the law to deal with externality problems. A person who is harmed by someone's actions can ask the courts to decide compensation. The court's decision will depend on whether or not he has a right to some good or service. Courts have established property rights for clean air, clean water, scenic views, sunshine, and quiet. If a person is not due compensation, then he does not have the property rights but the other party may have them. In this case he can pay the party harming him to stop the offending activity. However, victims of pollution often do not band together and sue the polluter, nor do they band together and pay him not to pollute. The difficulty with legal action is that there are serious problems (and thus large costs) in contracting and organizing large groups for legal action. One of these problems is the free-rider problem. Economist Ronald Coase pointed out that pollution problems would not exist if there were no difficulties and expenses in making contracts between polluter and victim. (We have mentioned Coase previously for his theory that the organized firm results from the difficulty of making many specific contracts.) The implication of Coase's work is that externalities should not be a small-group problem since legal remedies will exist if only a very few people are involved. It will only be a large-group problem.

Coase shows that property rights are not always a feasible way to solve the externality problem of "free" resources. Another solution is for the government to act as if it were the owner of these resources. The government does this when it regulates the number of ducks which hunters can kill. It says in effect that the government owns the ducks, and people cannot kill them without the permission of the government.

Government can charge for the use of its resources. It could, for example, charge polluters for the use of clean water and air. This charge would make polluters take into account the side effects of their activities (or in the jargon of economists, they would *internalize the externalities*), and would move the marginal cost curve in Figure 3 upward. There is some user fee (pollution tax) which would make the decisionmakers' marginal cost curves coincide with the marginal-cost-to-society curve, and thus correct the efficiency problem.

Government policy dealing with pollution and negative externalities has largely been one of regulation. For reasons that you can explore in question 5 at the end of the chapter, most economists believe that this is a less desirable (efficient) method of dealing with the problem than a policy of a pollution tax.

Finally there may not be a good solution to the problem of "free" resources for two reasons. First, the cost of a solution may be greater than the benefits of the solution. Most economists believe that there is some "optimal level" of pollution. Many productive processes produce waste products. These waste products, when considered damaging to people, are pollution. To remove them or to transform them into a form that no one considers damaging requires resources, and the use of those resources means that less of other products can be produced. Thus the reduction of pollution involves the weighing of costs and benefits as does virtually all other activity that economists discuss. The optimal level of pollution becomes that level at which the marginal benefit of any more reduction just equals the marginal cost of any more reduction. It is extremely unlikely that this level will ever be zero.

Second, there may be externality problems within the government just as there can be externality problems in the market. When there are externality problems in the market, we can call on the government as an outside agent to solve them. But if these problems exist in the government, there is no one to turn to. For example, suppose the citizens of a country are split into fifty special interest groups, and each group gets special benefits from the government. To pay for those benefits, the government must tax the citizens. The citizens end up paying a dollar in taxes to get eighty cents of special benefits. (Bureaucracy eats up the other 20%.) Though all would be better off getting rid of all special benefits, no one group will want to give up its special benefits, and the costs of organizing the fifty different groups to come up with an agreement may be very large. There may be no solution to this problem of the commons.

SUMMARY

This chapter examined the issue of income distribution, which is, along with the concept of economic efficiency, a main normative judgement used to evaluate government policies. The chapter looked at three different ways of evaluating fairness, and found problems with each. It then looked at possible problems the government would have in changing the distribution of income from that which the market gives.

The second topic of the chapter was the efficiency problems caused by resources which appear to be free, but which are in fact scarce. It showed that the problem of the commons was such a problem, and that it resulted in production inefficiency. It also looked at other sorts of externalities which resulted in inefficiency because the wrong mix of goods and services was produced. Finally, it suggested some ways to eliminate or minimize these efficiency losses.

QUESTIONS

1. Mehrgeld is an economy of price takers. The widget industry is unique in Mehrgeld because the production of widgets creates pollution. For each widget produced, the quality of air and water is lowered by $.50. This cost of $.50 is not shown in the table below.

Number of Widgets	Value to Buyers of Another Widget	Value of Resources Needed to Make Another Widget
1000	$10.00	$9.00
2000	9.50	9.00
3000	9.00	9.00
4000	8.50	9.00

a) Based on the table above, how many widgets will be produced by profit-maximizing producers?

b) What is the economically efficient number to produce?

c) What simple policy would an economist be most likely to recommend to internalize the externality of this industry?

2. The chapter said that there were three reasons that the earnings of a resource could change: the supply of the resource could change, the demand for the output which the resource produced could change, and the marginal productivity of the resource could change. Consider which of these factors would have changed for owners of land as a result of the Black Death, and predict what happened to rents.

3. In the past century both population and wages have risen. How can this be explained in terms of Figure 2?

4. At the end of Chapter 11 we considered a resource bundle with three uses. Use A had a value of $25 to people, use B a value of $22, and use C a value of $20. We said that a system of making choice which delivered the resources to use A was efficient, and argued that markets could be efficient. However, in Chapter 12 we saw a case (public goods) in which markets would allocate inefficiently. How will the resources be used in the cases below?

a) An industry which has the problem of the commons finds that when there are three producers, total revenue is $120. When there are four, total revenue increases by $20 to $140. Use C is adding the fourth producer.

b) An industry has positive externalities. If more is produced, buyers will get benefits of $21, but there will be spillover benefits of $4. This use of resources is use A.

c) An industry has negative externalities. If more is produced, buyers get benefits of $26, but there are spillover costs of $4. The extra production is use B.

5. Suppose you are the head of a pollution control agency and you desire to cut the pollution in Starry River by one half. The polluters, the amount they pollute, and their marginal costs of reducing wastes are given in the table below. Your staff has proposed three options. With Option A you require each firm to reduce wastes by one half. With Option B you require each firm to reduce wastes by five tons. With Option C you put a tax on the discharge of $4.50 per ton.

What is the cost to society (how much in resources is required) to clean up the river by one half under each option? What is the cost to business under each option? Which will economists be most likely to propose? Which will the Chamber of Commerce support?

Firm	Tons of Waste Per Day	Marginal cost of Reducing a Ton of Waste	Cost of A	Cost of B	Cost of C	
					Cleanup Cost	Tax Cost
Orion Ind.	10	$5.00				
Draco Co.	15	5.00				
Auriga Bros.	5	4.00				
Hercules Int.	5	6.00				
Pegasus Inc.	10	2.00				
Scorpio Entpr.	15	1.00				

(Comment: Economists consider taxes as a transfer which does not use resources. The notion of rent seeking suggests that assuming transfers as "free" is dangerous.)

6. Astronomers worry about "light pollution." How can there be such a thing?

7. Explain why people can consider the items below to be free. Then explain why (or when) they are in fact scarce. Explain the problem which occurs, and give a possible solution to eliminate it.

a. Ducks migrating south for the winter.
b. Travel on a freeway during rush hour.
c. Whales in the middle of the ocean.
d. Water in a river running through an industrial area.
e. Water in an aquifer in a fertile but dry area.
f. Buffalo in the 19th century.
g. Trees on public land on a heavily populated island that needs firewood.
h. Food in a college cafeteria for which students pay a fixed board fee.
i. Mammoths to ice-age hunters.

8. Should an employer take need into account when he determines how much to pay an employee? Should the government take need into account when it decides how much to tax a person? Most people answer "No" to the first question and "Yes" to the second. They believe that it is unfair for an employer to pay a man who has ten children more

than a single woman if they both do the same work, but it is acceptable for the government to tax the woman much more than the man. How can you explain this seeming inconsistency?

9. Economists believe that only humans matter. Suppose a waste product is harmful to animals but does not harm humans. Can it be pollution? Should we consider effects on animals and plants independent of effects on humans? Some people believe that plants and animals, as coinhabitors of the planet, have the same rights as humans. Do you agree or disagree?

CHAPTER 15

EFFICIENCY AND MONOPOLY

After you read this chapter and work the problems at the end, you should be able to:

Define duopoly, contestable market, barriers to entry.

Explain why a system with monopoly may be economically inefficient.

Explain what two problems a business must overcome if it is to price discriminate.

Explain why a monopolist which price discriminates may produce the economically efficient amount of output.

Explain why economics lacks a simple, complete theory of oligopoly.

Explain why economists argue that much regulation is sought by those regulated, not thrust upon them.

Whenever an economist recommends that a government take some policy action, he is no longer dealing only with positive economics. His policy recommendation must be based on some perception of what is desirable and what is undesirable, and this is a normative topic. Policy recommendations involve a merger of positive economics and normative judgements—a merger of understanding how the world works (which can, of course, be flawed) and belief about where it should be. In the normative realm the concept of economic efficiency is the most common criterion economists use. This chapter considers a topic which economists have discussed extensively using the concept of economic efficiency, the topic of monopoly.

WHAT IS WRONG WITH MONOPOLY?

Economists have rather consistently argued that monopoly is undesirable. On this issue they are probably not in disagreement with the public's rhetoric, though they sometimes see monopoly problems in different places than the public does.

A monopoly is an industry in which there is one seller. Because it is the only seller, the monopolist faces a downward-sloping demand curve, the industry demand curve. The downward-sloping demand curve means that if the monopolist wants to sell more, it must lower its price. (We are assuming that price discrimination is not possible, that the firm can charge only one price.) Because the monopolist must lower price to sell more, the extra or marginal revenue it gets from selling more is less than the price it charges. Thus its marginal revenue curve lies below its demand curve. In contrast, for a seller who is a price taker, demand is identical with marginal revenue, as Figure 1 of Chapter 11 illustrates.

Table 1 illustrates the case of monopoly. Marginal cost is the cost (or value) of the additional resources needed to produce another unit of output. The marginal benefit to consumers is the price which consumers are willing to pay for each unit. You should recognize this column in Table 1 as a demand curve. The maximization principle tells us that the economically efficient amount to produce is five. This amount gives consumers the greatest value. To produce the first unit, the firm takes resources which have a value of $5.00 and turns them into something with a value of $7.01. Since this transformation has increased value, producing the first unit is more economically efficient than producing none. By this logic, producing the sixth unit would decrease economic efficiency since in doing so the firm would take resources with a value of $5.00 and transform them into something with a value of only $4.51.

The monopolist, however, will find it most profitable to produce only three units. It does

TABLE 1
THE SIMPLE ANALYTICS OF MONOPOLY

Output	Marginal Cost	Marginal Benefit to Buyers	Marginal Benefit to Seller
1	$5.00	$7.01	$7.01
2	5.00	6.51	6.01
3	5.00	6.01	5.01
4	5.00	5.51	4.01
5	5.00	5.01	3.01
6	5.00	4.51	2.01

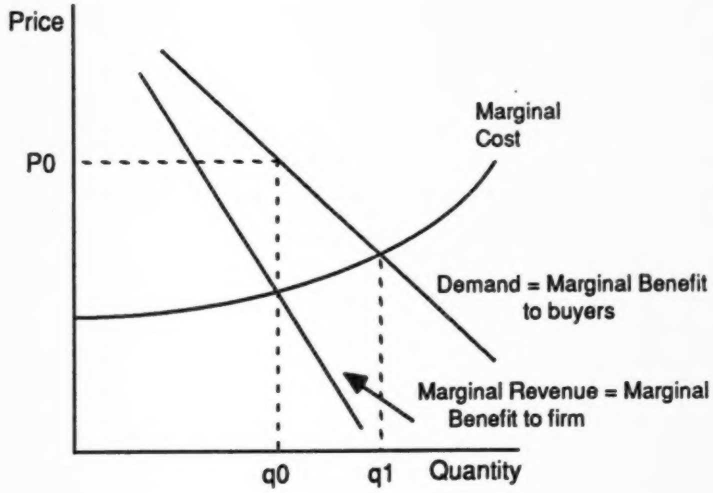


Figure 15.1 Monopoly dissected.

not see marginal benefit the same way as buyers see it. For the seller, the extra benefit of the second unit is only $6.01. It sells the second unit for $6.51, but to sell the second unit, it had to reduce the price it charged by $.50. Thus it "lost" $.50 on the first unit, so the net increase in its revenue was only $6.01. (You can get the same answer by computing a total revenue column and then calculating revenue increases.) In a similar manner, the rest of the fourth column can be obtained. Using the maximization principle, one can see that producing beyond the third unit is not in the interests of the firm. The fourth unit brings in added benefits of only $4.01 to the firm (it sells for $5.51, but to sell it, the firm lowers price by $.50 on three other units), but costs an added $5.00. From the point of view of the buyers, however, the fourth unit should be produced. It brings them added benefits of $5.51 and uses resources worth only $5.00.

The discussion above can be illustrated with Figure 1. The seller attempts to set marginal cost equal to marginal revenue, or to produce at q0. From the consumers' viewpoint, the best amount to produce would be q1. The monopolist restricts output because of a divergence between marginal benefit as the firm perceives it and marginal benefit as buyers perceive it. Producing beyond q0 is not in the interests of the firm because the extra benefit it sees, the marginal revenue curve, is less than the extra cost of production, shown by the marginal cost curve. Extra output is in the interests of buyers because the extra benefit they get, shown by the demand curve, is greater than the extra costs of production.

In terms of a production-possibilities curve in Figure 2, an economy with some industries competitive (all transactors are price takers) and others monopolized (sellers are price makers) will produce at point *b*. However, consumers would be better off at point *a* because the gain of *x* amount of monopolized goods has a greater value than the loss of *y* amount of competitive goods. Because a monopolist restricts production from what a competitive industry would do, too many resources are being used in the competitive industry and not enough in the monopolized industry. In terms of the three kinds of efficiencies necessary for economic efficiency

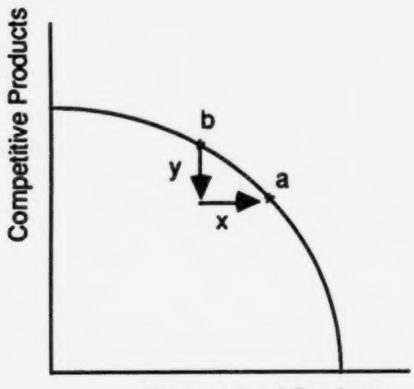


Figure 15.2 Monopoly restricts output relative to competition.

which Chapter 11 discussed, the existence of monopoly violates the third. Marginal rates of substitution are not equal to marginal rates of transformation.

When one understands that monopoly leaves unexploited value, one can ask two questions. First, one can ask whether people have found ways to capture this value. Since the search for value occupies a great deal of people's talents and energies in a market economy, it should not surprise us to learn that there is a way to capture this value and it is extensively used. Sellers can capture it with price discrimination. Second, one can ask if the government can eliminate the unexploited value with some sort of intervention. Economists have suggested two important ways for the government to intervene, through antitrust actions and through regulation.

However, before we turn to these topics of price discrimination, antitrust, and regulation, we look at situations which may be near-monopolies. In these situations there is more than one seller, but the sellers may act like a monopoly by restricting transactions below the economically efficient level and creating an unexploited value.

THE THEORY OF FEW SELLERS

We have just seen the theory of a market with one seller and many buyers. In Chapter 11 we saw the theory of a market with many sellers and many buyers. What changes in these theories would be needed to tell us what happens if there is more than one seller, but less than many? What, for example, happens if there are two sellers, or **duopoly**? Though many economists suspect that the results of two sellers are more similar to those of one seller than to those of many sellers, economics lacks a clear theory that can prove this suspicion.

The problem of **interdependence** has thwarted economists' attempts to develop a good theory of oligopoly. When there are only a few sellers, each recognizes that his decisions affect not only himself, but others, and that the others may react to what he does.

The problem of interdependence can be shown in terms of game theory in a situation that is identical to the prisoners' dilemma of Chapter 2. Table 2 shows the pricing options of the only two gas stations in an isolated town. The payoffs in the center are profits. If both stations charge high prices, the joint profits are maximized. But then each has a temptation to cut prices to get to a more favorable corner. If one of them gives in to this temptation, he may start a gas price.war as the other firm must retaliate. They then end in the least favorable position of lowest joint profits. We could expand the table to show what would happen if they cut prices even further, to $.80, and we would see that this cut would make them both still worse off. Gas price wars do break out occasionally, but they usually do not last long. Rather quickly the firms realize the futility of the fight, and will try to return to the original position.

There may be no equilibrium solution in a situation of this sort. Rather there may be a period of *collusion* in which firms agree (though it may be an unspoken agreement) to

TABLE 2
THE PRISONER'S DILEMMA AND DUOPOLY

PRICING STRATEGIES		SAM'S SUPER SERVICE
	$1.00	$.90
$1.00	100 100	130 10
$.90	10 130	20 20
GUS' GAS		

keep prices high, then the collusion may disintegrate as firms begin cheating, and finally a new period of collusion may begin. Whether sellers collude or compete will depend on many factors which can be difficult to measure and put into a theory, such as the number of sellers, their personalities, whether they have equal or unequal shares of the market, whether their costs are the same, the ease of cheating and of detecting cheating, and whether the sellers can compete on nonprice bases such as with service and quality.

A thorough examination of possibilities of oligopolistic strategies and how well they fit observed behavior of real-world oligopolies is a large and controversial subject beyond the scope of this book. The point that is important in dealing with efficiency is that if oligopolists perceive their demand curves as downward-sloping, that is, if they take into account that the amount they produce will have a significant effect on the price they can charge, then their marginal revenue curves will lie below their demand curves and they will restrict output relative to what an industry of price takers would. Thus there will be an efficiency loss involved. Most economists use the term "market power" to describe the ability of any price maker to set price.

Though the word "power" usually implies the ability to use force, the possessor of market power usually does not possess this ability. A monopolist cannot usually fine, imprison, or physically harm someone who does not do what the monopolist wants. However, the term "market power" is widely used and there seems to be no good alternative. Hence, we will use it below.

When the possession of market power is profitable, it should attract new entrants into the industry. If entry is easy, then the existence of very few or even only one firm may not result in economic inefficiency. The threat of potential entry may be enough competition to keep the industry operating at or close to the competitive solution. In this case the market is a **contestable** market. However, if entry is not easy, but there are significant **barriers to entry**, the threat of competition is less. Barriers to entry exist when there are **sunk** costs, expenses which cannot be recovered once a firm has entered the industry. Where these costs are high, the industry probably operates as the theory of monopoly suggests it will.

Barriers to entry can take several forms. They will exist if large amounts of specialized machinery is required to enter an industry, and resale of that machinery is difficult. They will exist if a firm must establish a reputation for the quality or reliability of a product. They will also exist if a firm must expend resources in order to get governmental approval to enter. In all of these cases, the barriers to entry can be viewed as sunk costs.

Some economists also treat situations in which there is rent or producer surplus as barriers to entry. Rent exists when some resource is especially good for a particular use. Land of special fertility, with rare ores, or with unusually good location are difficult to compete against, as are people with special talents.

The theory of contestable markets suggests that even if there is only one seller, the seller may be forced to act as if there were many more. In contrast, there are times when great numbers of sellers are able to organize and act as a unified seller. Sellers have the incentive to act in this way because it will increase profits. The key to their success is their ability to restrict sales.

Important cases of such restrictions occur in the resource market. Unionization is one way that sellers in the resource market can restrict sales to raise price. A union can negotiate a contract above the wage the market would set. This possibility is shown in Figure 3 where Pn is the negotiated wage and Pm is the market wage. The firm agrees to this because the alternative is a strike. If the firm could simply replace striking workers with nonstriking workers (called **scabs** by unionized workers), the threat of a strike would have much less effect. But attempting to do this is likely to cause violence and destruction. The surplus indicated on the graph should show up as a large number of applicants whenever the firm hires. Rationing can take place on a first-come-first-served basis, on the basis of "who you know," or any other basis the firm and union decide on.

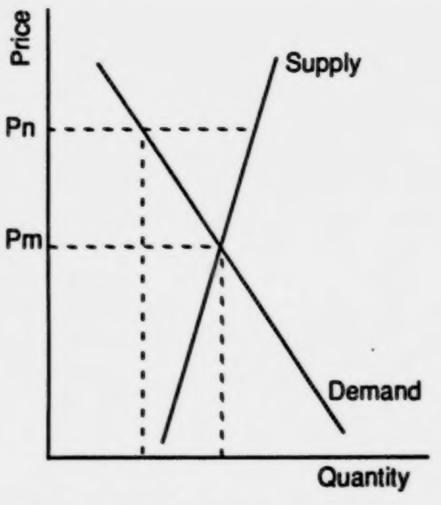


Figure 15.3 One way to raise wages is to ration jobs.

An alternative way to restrict transactions is to restrict supply. This option usually requires the cooperation of the state. One way it can be done is with licensing. If one cannot enter a profession without a license, and if the number of licenses can be restricted, then fewer people will enter the occupation. This method is shown in Figure 4 where *s1* is the supply curve without licensing and *s2* is the supply curve with it. Among occupations that use this method (the list varies from state to state) are teachers, most medical professions, lawyers, accountants, beauticians, barbers, morticians, and veterinarians.

For the use of resources to be efficient, their prices must reflect their value that they could contribute elsewhere in the economy. The price which a person using a resource sees should be the opportunity cost of that resource. The problem which licensing and union power cause is that they generate prices which misstate the opportunity costs of resources. When the cost of a resource which a decision maker uses differs from its social cost (which is its opportunity cost), decisions will be distorted and not economically efficient. For example, if a union contract says that an employer must pay Ian Segoe $10.00 per hour if he works for it, and Ian will only add value of $9.00, the firm will not hire him. Instead he must take a job that pays only $8.00 per hour. Since Ian's labor is not used where it adds the most value, the economy is economically inefficient.

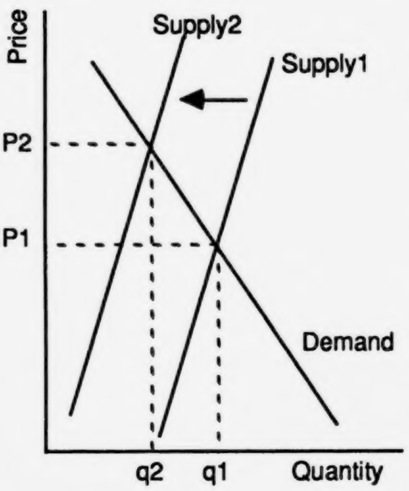


Figure 15.4 Another way to raise wages is to reduce supply.

One might at this point object to treating unions and professional licensing in the same way that monopolies are treated. (But recall the opening of this section, which says economists see monopoly problems in places the public does not.) Many people believe that labor unions aid working people and that licensing protects the public. They believe that the extra wages unions obtain for their members come from the profits of big business, and without licensing the public would be exposed to charlatans. There is often some truth in these beliefs, but not the whole truth.

Effective unions do raise wages of their members. All economists who have studied the data seem to agree on this point, though they differ on how much unions raise wages of their members relative to those not in unions. Estimates range from 5% to 30%. These studies also suggest that the increase in wages for union members has the effect of pushing down wages in nonunionized sectors. Figure 4 shows that the number of workers employed in an industry with successful unions will decrease, and thus there will be more workers available in nonunionized sectors, causing the supply curve to shift to the right. The equilibrium wage rate in the nonunionized sector will be lower as a result of unions.

This is not to say that economists see no useful role for unions. In a modern economy market relations have been partially replaced by the hierarchical relations within the large corporation. There is a potential for abuse in the power relations which exist within the firm, and unions are a way to respond to this. Unions are more likely to be formed in organizations which have poor management-worker relations than in those where management-worker relations are harmonious.

Licensing may improve the quality of services, but at the expense of reducing quantity. Hence, even if licensing achieves its stated purpose, of improving quality, it is not necessarily true that the consumer is better off. The loss of cheap, low-quality service may offset the higher quality; would automobile drivers be better off if they could only buy Rolls Royces? However, there is also considerable doubt on the part of some economists, based on their reading of the data, that licensing does much to improve quality.

PRICE DISCRIMINATION

When a family with small children goes to a restaurant, they are often given a children's menu in addition to the regular menu. If they order two similar items, one from each menu, they will find that the item ordered from the children's menu will be a bit smaller, but its price will be *much* smaller. In fact, it would often be worthwhile for the entire family to order from the children's menu, but they cannot. Restaurants only allow children to order from it.

Why do restaurants use children's menus? Perhaps restaurant owners love children. A more convincing argument would be that for some reason they find that offering children's menus is profitable. It can be profitable if adults who come to restaurants with children are, on the average, more sensitive to prices on menus than adults who come to restaurants without children.

Children often do not appreciate restaurant food and service, and often waste a large part of their food. Parents know this, and do not want to pay a lot for their child's meal. If restaurants treat children like adults, the restaurants may lose customers as families switch to fast-food restaurants. If this explanation is correct, then restaurants price discriminate.

A seller price discriminates when it charges different prices to different buyers. The ideal form of price discrimination, from the seller's point of view, is to charge each buyer the maximum which the buyer is willing to pay. If the seller in Table 1 could do this, it could charge the first buyer $7.01, the second buyer $6.51, etc. In this case the marginal revenue curve becomes identical with the demand curve. The seller will sell the economically efficient amount, it would capture the entire consumers' surplus, and it would substantially increase profits.

Every seller would price discriminate if there were not two major obstacles standing in the way. First, the seller must be able to distinguish between those buyers who are willing to pay a high price from those who are not. Second, there must be substantial difficulty for a low-price buyer to resell to those willing to buy at a high price. (A third obstacle standing in the way in the United States is legal. Several laws restrict the ways in which businesses can price discriminate. Though the stated purpose of these laws is to preserve competition, the sponsors clearly wanted to prevent competition from eliminating small sellers with high costs.)

Because price discrimination is potentially profitable, businesses have found many ways to do it. Theaters often charge younger customers less than adults. Doctors sometimes charge the rich or insured patient more for services than they charge the poor or uninsured. Grocery stores have a lower price for people who bother to check the newspaper and clip coupons. Some companies, such as firms selling alcoholic beverages, produce similar products but try to promote one as a prestige brand with a much higher price. Electric utilities usually charge lower rates to people who use a lot of electricity (an thus probably have electric stoves and water heaters) than they do to those who use only a little electricity (and who probably have gas stoves and water heaters).

To the extent that businesses find ways to price discriminate, they eliminate the triangle of welfare loss in Figure 1 and approach the economically efficient amount of production. Thus the mere existence of monopoly does not prove that there is economic inefficiency.

ANTITRUST POLICY

Governments have a number of policies that affect monopoly and market power. Two which are intended to reduce monopoly are regulation and antitrust policy. However, there are also policies which purposely increase monopoly. Patent laws, for example, encourage market power. The belief that it is socially useful to encourage people to develop new processes and products has led to laws which provide monopoly rewards for those who do so. Patent laws are a case in which there is some recognition that the analysis of efficiency is static but the economy is dynamic.

Competition may mean price takers in a static model, but in a dynamic model it can be striving for positions of temporary monopoly. Joseph Schumpeter argued that this competition for positions of temporary monopoly was the most important type of competition in a market economy. He argued that it provide a *gale of creative destruction*. Entrepreneurs would find new products and ways of producing things which would make obsolete the old products and ways of production. In turn their positions would sooner or later be destroyed by new entrepreneurs. For Schumpeter the essence of market economies was change.

Despite Schumpeter's argument, most economists have argued that the government should take measures to attack the efficiency loss of monopoly. **Antitrust policy** is one way to do this. Antitrust policy intends to make companies act in a competitive manner. It does this by breaking up companies that are monopolies, by prohibiting mergers which would increase market power, and by finding and fining companies which collude to establish higher prices.

The antitrust laws in the United States are stricter and more comprehensive than those of other industrialized nations. Many other industrialized nations, perhaps because they focus on the possibility of foreign competition, have minimal antitrust law. In the United States the original antitrust law was the Sherman Act of 1890. Important extensions of the law were passed in 1914.

In the United States two government agencies have authority to enforce antitrust laws, the Antitrust Division of the Justice Department and the Federal Trade Commission. In addition, the laws allow companies, organizations, or persons who are adversely affected by actions which violate the antitrust acts to sue in court for damages.

The earlier part of this chapter and Chapter 11 built the economic case for antitrust policy. Monopoly can lead to an inefficient use of resources when compared to the competitive result. Furthermore, there are clear cases in which businesses have tried to act as monopolists, charging high prices with restricted outputs.

Nonetheless, there are unanswered questions about how great the net benefits of antitrust enforcement are. There are conditions under which monopoly may be more efficient than competition. When an industry has increasing returns to scale, two small firms will require more resources to produce a given amount of output than one large firm. Trying to keep two such firms in existence by, for example, preventing merger will keep the economy inside its production-possibilities frontier. The economy will be production inefficient.

Further, the argument that competition is more efficient than monopoly relies on static, equilibrium analysis which assumes fixed production functions, demand curves, and supply curves for resources. Some economists, following the lead of Joseph Schumpeter, argue that the competition that matters in a market economy is the struggle to find ways to change these constraints and to gain temporary monopoly power.

It is possible that the laws can be used to attack behavior which is not only not economically inefficient, but which may be desirable. For example, suppose a company is extremely well-managed, excelling at producing quality products at low prices. Such a company will be profitable and will grow to dominate its industry. Its dominance and large profits would appear to indicate that it was a monopoly and could trigger antitrust action. In fact, there are actual cases similar to this. The Alcoa case of 1945 seems to be one.

Alcoa was found innocent of monopolistic practices by the district court, but an appeals court reversed this decision and found it guilty. Justice Learned Hand explained in the decision:

> "It was not inevitable that it [Alcoa] should always anticipate increases in the demand for ingot and be prepared to supply them. Nothing compelled it to keep doubling and redoubling its capacity before others entered the field. It insists that it never excluded competitors; but we can think of no more effective exclusion than progressively to embrace each new opportunity as it opened, and to face every newcomer with new capacity already geared into a great organization, having the advantage of experience, trade connections and the elite of personnel."

Hand's decision makes clear that Alcoa maintained its position with managerial excellence. There was no evidence that it restricted output to keep prices high, the sins of monopoly which lead to economic inefficiency. On the contrary, it kept prices low and expanded output. Hand's reasoning, subsequently reaffirmed in other cases, is difficult to justify economically.

There are clearly cases in which antitrust laws move the economy closer to a competitive state. There also seems to have been times when antitrust actions were misguided, when antitrust actions did not result in lower prices and higher quantities for consumers. Unfortunately economists have traditionally been so sure that antitrust laws were beneficial that they have not sought to measure whether the net benefit of those laws was in fact positive. We really do not know how much good, if any, antitrust actions accomplish.

REGULATION

If increasing returns to scale cause monopoly to form, antitrust policy may not be appropriate. With increasing returns, one large firm can produce at lower costs than several smaller firms. Monopolies which form for this reason are often called **natural monopolies**. Breaking them up into smaller firms makes no economic sense because it increases costs, and perhaps price.

In these cases economists have argued that regulation may be appropriate. Government regulation involves putting limits on what a business (or consumer) can do. Certain activities become illegal and others become mandatory. Regulations can take many forms. They can forbid or require certain prices, or certain production techniques, or certain products.

The ideal form of regulation for a monopoly would be to force it to set its price equal to its marginal cost. This would force it to price as it were a price taker. The monopoly should find it profitable to increase output to the economically efficient level. (See questions 1 and 2 at the end of the chapter.) Life, however, is more complicated than theory.

The publicly-stated goal of almost all regulation is to improve the well-being of the public. If economists were convinced that publicly-stated goals were achieved, they would not find the topic of regulation very interesting. But their private-interest view of government, discussed in Chapter 13, gives a rather different view of how regulation works than the view with which most people are familiar. In the private-interest view regulation only comes into being when it is in the private interests of the legislators to bring it into existence.

Legislators are elected in a democracy, and it is in their interests to do those things which contribute to their re-election. (Some legislators come from districts which are relatively safe, and these legislators may have more leeway in voting their prejudices, ideology, or whims than others who are in less safe districts.) To be re-elected, legislators need votes, and to obtain votes, they need campaign funds and volunteer workers.

Groups which seek regulation must be able to supply either votes or money or both. Because there is a cost to getting regulation, a group will not seek it unless the issue is of considerable importance to the group. It is difficult to organize groups that are very large and in which each individual has only a small stake in the central issue. This difficulty arises because of the free-rider effect. Also, if an issue is of only small importance to an individual, it is unlikely that he will vote on the basis of a candidate's stance on that one issue because there are so many different issues on which most candidates take a stance.

One group that can often be organized is a group that is or will be regulated. The effects regulation has on them will be very visible to this group, and the group is usually relatively small relative to other groups which are affected by the regulation. As a result, some economists have proposed that regulation often comes about because it is sought by the very groups which are regulated. The most prominent proponent of this view has been George Stigler.

This idea may at first seem to contradict common sense. Regulation involves a reduction in choice, and thus should make those regulated worse off. However, recall the model of the prisoners' dilemma. In a case of this sort, decision makers will be willing to have their choices restricted provided that all others have their choices restricted in the same way. Thus if regulation can increase prices by restricting competition, either through a minimum price or control of entry and substitutes, there are potential gains for sellers. There are risks and costs involved in obtaining regulation, however. First there are costs of organizing and lobbying. If a group is large there may be free-rider problems (though the potential for unfavorable treatment if the regulation is passed may help offset this). Then once regulation is in place, it may not work precisely the way that those who sought it thought it would. Outsiders will be involved in the industry, and there is the possibility that they may be unfriendly. Also, regulation involves red tape; this will be a cost to those who are regulated.

The interest group will be successful in establishing regulation if it can promise the legislators substantially more votes or money (usually campaign funds to obtain more votes, but not always) than can the opponents of the regulation. The reason it will be able to control the regulatory agency, or "capture" it, involves the behavior of a different part of government, the bureaucrats.

The economic hypothesis assumes that the bureaucrats—non-elected government officials—are motivated on the basis of narrow self interest. (Did you expect any other assumption?) They serve in an organized group, their agency or department, and in many ways their behavior will be shaped by the constraints the agency faces just as the behavior of those in a corporate business is shaped by the constraints the business faces. The regulatory agency must purchase resources, including services of lawyers, accountants, and economists. It must produce an output, which is the regulation. Finally, it must obtain funds for its continued existence. The funds it obtains depend ultimately on its product, regulations. If the regulations are made sloppily, they may be challenged in the courts and overturned. If this happens too frequently, the agency may attract unfavorable attention and this may result in funds being cut. Or if its regulations unfavorably affect organized special-interest groups, these groups will through lobbying efforts try to cut the funds that the agency is given or try to change the personnel who run the agency.

The bureaucrats who make up the regulatory agency have a variety of goals, just as those who make up a corporation have a variety of goals. In both cases some of those goals can be realized only if the agency survives. Bureaucrats who threaten the survival of their organization will face dismissal if that can be done, or strong pressures to change behavior. Unlike the customers of a business firm, who can "leave" and take their business elsewhere if the product does not satisfy them, "customers" of a regulatory agency cannot leave. But they can voice their complaints and in general try to make life miserable for those responsible.

The above summary outlines the approach that some economists have taken toward the issue of regulation. This approach is useful only if it is consistent with facts that are available. Can cases be found where those regulated sought regulation? What evidence is there that those who are regulated want to be regulated? Only if the answers to these questions suggest that the economic approach has some validity is it worthwhile to discuss what it implies for economic efficiency.

Economic historians and students of regulation have found substantial evidence that the railroads supported the establishment of the Interstate Commerce Commission, that utility companies supported the establishment of state utility regulation, and that various occupational groups have usually been the main proponents of occupational licensure. Further there have been a number of studies which have shown that occupational groups earn more in states in which they are regulated than in states in which they are not. Finally, those groups which are regulated often oppose most vigorously efforts to repeal regulation. However, there are some regulations which have not been supported by those who are regulated. Regulations which affect many different industries, such as regulation intended to promote worker safety, equality of employment, and pollution cleanup, have not generally been sought by groups which are regulated. Nonetheless, there is considerable evidence that much regulation fits more closely the private-interest hypothesis than the public-interest hypothesis.

The private-interest theory of regulation suggests that much regulation will create economic inefficiency. A considerable amount of regulation involves industries which have a substantial amount of competition (such as agriculture, transportation, and oil and gas production) and both the theoretical and empirical evidence suggests that this regulation will lead to higher prices and lower quantities, which is economically inefficient. The adverse effects of regulation can be partially offset by competition in service if competition is not allowed in price. Thus airlines, which could compete only to a limited extent on price during the years in which they were tightly regulated, could and did compete on the quality and quantity of service. They advertised the quality of meals, the frequency of flights, the courtesy of stewardesses, and the availability of movies.

Paradoxically, regulation of competitive industries may in the long run offer no financial benefit to those regulated, but repealing it may cause considerable harm to the industry. For example, taxi service is regulated in New York City, and the number of cabs is limited. To operate a cab one must have a permit to operate, called a medallion. Someone entering the industry must buy a medallion from someone who already owns one. The value of this medallion reflects capitalized value of the greater than competitive returns that regulation brings. A new entrant thus finds only a competitive return in the industry once the cost of the medallion is taken into account. (In recent years medallions have cost more than $50,000.) If regulation ceased, those taxi owners now in the market would suffer a capital loss since their medallions would cease to have value. The only taxi owners who seem to have truly benefitted from regulation were those in business at the time regulation began. They were able to capture the full benefit of the regulation, leaving none for those who entered the business later.

Though theory says that regulation of monopoly can increase economic efficiency, empirical studies have not yet shown that this regulation actually increase efficiency in practice. Ideal regulation from an economic efficiency standpoint makes a monopoly set price equal to marginal cost. However, marginal cost figures only make sense if the firm is minimizing its costs, and once the government starts trying to lower price to marginal cost, the firm has an incentive to increase costs. As a result, government almost never tries to set price equal to marginal cost.

The alternative which is usually used is for regulators to allow a monopoly only a competitive return on its investment. This type of regulation gives the firm two incentives, neither socially desirable: to increase costs (through higher salaries and perquisites, for example) and to invest more than they need to. Regulation almost always contains provisions protecting the regulated monopolist from any competition. Theoretically one can show that under certain conditions the public is served when monopoly is protected, but whether the conditions needed to make this argument hold exist in the case of any regulated monopolies is uncertain.

A strong case for nonprice regulation (such as environmental, safety, and health) can also be made on theoretical grounds, but again there is doubt as to the extent that actual regulation lives up to its theoretical potential. Part of the problem is that lack of good data with which to test hypotheses about the effects of this regulation. Economists have been unwilling to accept regulation as desirable simply because its publicly stated goals are desirable. They have tended to argue that the regulation will have desirable effects only if the proper incentives are created.

An example of a case in which incentives may not be ideal is the program of approving new drugs that the Food and Drug Administration runs. To be approved for use in the United States, a new drug must undergo a series of time-consuming and expensive tests to show that it is safe. This procedure means that unsafe drugs are unlikely to be allowed onto the market. An example was the drug

Thalidomide which was released in Europe but which never approved for use in the United States. Thalidomide caused serious deformities in infants when taken by mothers during early stages of pregnancy. But this procedure also means that safe drugs which could save lives will be delayed (and hence lives lost), and drug companies have little incentive to develop drugs for rare diseases because the cost of testing diminishes the probability of ever making a profit. This structure of approval is almost inevitable. Deaths and problems which an unsafe drug causes are highly visible and widely publicized. They will cause unfavorable publicity and embarrassment to the agency, including perhaps Congressional inquiries. Deaths caused by the unavailability of drugs are far less noticeable, and do not attract the same intensity of publicity.

The subject of regulation is a changing and developing field, and all conclusions it has reached are subject to change. The basic theory sketched in this chapter only scratches the surface of this subject. Regulation is subject to a variety of influences besides the interests of those who are regulated, and these interests can modify regulation in a variety of ways. Further, when a regulatory agency is securely established, it can have considerable leeway in opposing the interests of those who are regulated.

The most important point that emerges from economic study of regulation is that the publicly-stated rationale for government action may differ from its true purpose, and that intended results may differ from actual results. Arguments which ignore these distinctions may have wide appeal to the general public, but they will not be taken seriously in scholarly discussion.

SUMMARY

This chapter showed how the concept of economic efficiency shapes the way economists approach policy decisions concerning monopoly. It showed that the introduction of monopoly into a competitive economy results in economic inefficiency because the monopolist sees the marginal benefit of its actions in a different way than consumers do. Monopoly is inefficient because it prevents an economy from producing the mix of products which have the greatest possible value.

The chapter also looked at the theory of oligopoly. It discussed how businesses can price discriminate and why this may eliminate the welfare loss due to monopoly. It also examined two ways that the government can try to correct the problem of monopoly, antitrust policy and regulation, and it discussed some of the second thoughts economists have had about these policies.

QUESTIONS

1. Assume that a firm faces the marginal cost schedule and the demand schedule below. Also assume that the firm will maximize profits.

Output	Marg. Cost	Price Consumers will pay
1	$6.90	$10.00
2	7.00	9.50
3	7.10	9.00
4	7.20	8.50
5	7.30	8.00
6	7.40	7.50
7	7.50	7.00
8	7.60	6.50

a) What price would a monopoly charge and how much would it produce if it can charge only one price? (Hint: it helps a lot to compute marginal revenue.)

b) How much would the monopolist produce if it were a perfect price discriminator?

c) How much will the monopolist produce if a price ceiling is introduced at a price of $7.50? (Hint: See question 5 at the end of Chapter 10.)

2. In the question above, (a) What is the size of the producers' surplus with a monopoly price? What is the size of the consumers' surplus? What is their total?

b) What is the size of the producers' surplus with perfect price discrimination? What is the size of the consumers' surplus? What is their total?

c) What is the size of the producers' surplus with this price ceiling? What is the size of the consumers' surplus? What is their total?

d) Which of the above three situations will be economically efficient? Explain. Can the consumers prefer an inefficient situation to an efficient one? Explain.

e) Suppose that the monopolist is a polluter, and that for each unit of output it produces, it imposes a burden on the community of $1.50. What does this do to the economically efficient amount of production?

f) Suppose the government, in an effort to increase economic efficiency, imposes the $7.50 price ceiling on this monopoly which is also a polluter. Will the policy be successful? Explain. (Comment: In a world full of imperfection, eliminating one may or may not improve the situation.)

3. In order to price discriminate, a seller must be able to separate its market into two parts. One part must contain those who are sensitive to price, and the other those who are not. Explain how each of the following strategies separates buyers into these two groups:

a) A home computer maker offers a $50.00 discount to any buyer who trades in another home computer or video game.

b) Movie theaters charge a reduced admission for children under 12.

c) Grocery stores charge lower prices on some items if people have coupons which they clipped from the paper.

d) Colleges which are prestigious and expensive give scholarships on the basis of need.

e) Colleges which are not prestigious and not expensive give scholarships on the basis of academic ability.

f) Economics journals charge libraries higher subscription rates than they charge individuals.

g) The American Economic Association has membership rates which are higher for those who earn large incomes than for those who earn small incomes.

4. At the end of Chapter 11 we considered a resource bundle with three uses. Use A had a value of $25 to people, use B a value of $22, and use C a value of $20. We said that a system of making choice which delivered the resources to use A was efficient, and argued that markets could be efficient. However, in Chapters 12 and 14 we saw cases in which markets would allocate inefficiently. How will the resources be used in the cases below?

a) A firm is a monopoly. It can sell four for $26 and a fifth for $25. Producing and selling the fifth is use A.

b) A firm with use A is a monopsonist. The resource bundle costs $20, but it will be the tenth bundle the firm has bought. When it buys 9, it pays $19 per bundle. (Use B does not involve monopsony.)

5. It is time to return again to Scimonoce. In the previous chapter you saw what happened if there were numerous sellers, each trying to maximize profits and each unaware that his decisions had an influence on price. Below you will see what happens if there is but one seller, which means there is monopoly.

a) Suppose that one person owns all the farm land. What will the total revenue of the farms be if 50 units of corn are produced? (Hint: $1.00 multiplied by the amount of oats plus the price of corn multiplied by the amount of corn produced. See your answers to Questions 6 and 7 in Chapter 11).

b) Suppose the monopolist decides to produce 90 corn. What is the total revenue he could obtain from all five plots of land?

c) Find how much he could obtain from no corn, 50 corn, 120 corn, 140 corn, and all corn. Which level of output is best for the monopolist? Is it economically efficient?

d) Why does a competitive system lead to a different result than a monopoly? (Hint: Does a competitive firm worry about the impact of a change in quantity on price? Do you see a free-rider effect here? Does a monopoly worry about the effect of changing quantity on price?)

6. A seller faces two markets. At home it faces the demand curve shown in column 3 below. If it sells one unit, it can charge $16, if it sells a second, it must lower price to $15, etc. In the foreign market it is a price taker and can sell as much as it wants at a price of $8.50.

a. If the seller could only sell in the home market, how many should it sell to maximize profit? (Hint: Solve by completing column 5, filling in the marginal revenue from selling in the home market. Then compare this column to marginal cost in column 2.) What price should the seller charge? (Hint: Use the demand curve in columns 1 and 3.)

1 OUTPUT	2 MC	3 P1	4 P2	5 MR1	6 MR2	7 MRtot
0						
	7.00					
1		16.00	8.50			
	7.20					
2		15.00	8.50			
	7.40					
3		14.00	8.50			
	7.60					
4		13.00	8.50			
	7.80					
5		12.00	8.50			
	8.00					
6		11.00	8.50			
	8.20					
7		10.00	8.50			
	8.40					
8		9.00	8.50			
	8.60					
9		8.00	8.50			
	8.80					
10		7.00	8.50			

b. If the seller can sell only in the foreign market, how many should it sell?

c. Suppose the seller can sell in both markets, charging $8.50 in the foreign market and any price it wants in the home market. How many should it sell in the home market and for what price? How many should it sell in the foreign market? (Hint: Solve by computing a marginal revenue column for both markets. Assume that the seller sells the next unit in the market in which it obtains the most marginal revenue. As long as marginal revenue exceed $8.50, it will sell in the home market, and thereafter it will sell in the foreign market. Compare this column to the marginal cost column to obtain total sales.)

d. How much more profit does the firm make when it can price discriminate in part c than in part *b*?

7. A firm can produce 1,000,000 light bulbs at a total cost of $1,000,000. It sells them for $1.10 each, making a profit of ten cents per bulb. One of the executives finds that if they produce two million bulbs, total cost rises to $1,500,000. This executive also finds that the company can sell these light bulbs abroad for $.65 each. He argues that the firm

should expand production and sell the extra bulbs abroad. Other executives argue that this plan is nonsense. They point out that if the firm produces two million bulbs, the average cost per bulb is $.75. They state that no company can make a profit by selling its product for less than the cost of production.

Who is right in this case? Explain.

8. "Monopoly is bad and competition is good" is a simple way of summarizing the basic result of this chapter. However, one must be careful with this conclusion, because under some conditions it is false.

Suppose for example, that we have an industry with increasing returns to scale. Small plants have a constant marginal cost of $2.20 per unit. Large plants have a constant marginal cost of $2.00, but have a fixed cost of $12.00. The industry has the following demand schedule:

Amount	Price	Elasticity of demand	Marg. Rev
50	$2.30	23	
100	2.20	11	
150	2.10	7	
200	2.00	5	

a) Show that for output below 60 the small plant has lower costs, and that for output above 60 the large plant has lower costs.

b) Based on the marginal cost data and this demand curve, what is the economically efficiency amount of output?

c) If the industry is a profit-maximizing monopoly, how much output will it produce? Why? What will its profit be? (Hint: Use elasticity and price to compute marginal revenue. See equation 10 in Chapter 9.)

d) Which of the three types of economic inefficiency discussed in Chapter 11 is involved here?

e) Suppose that an aggressive antitrust action breaks up this monopoly into four equally sized firms. Further, suppose that each of these firms acts as if it were a price taker. This assumption makes the industry perfectly competitive. At what level of output will the industry operate when it is competitive. (Hint: What plant size is best for small levels of production? A price taker sets MC equal to what? What will the industry output be? Each firm will produce one-fourth of it.)

f) If competition is better than monopoly, the inefficiency should disappear. Does it? What type of economic inefficiency discussed in Chapter 11 is involved in part e? (Hint: compare the value of resources used with monopoly as opposed to competition.)

g) Is the monopolistic or competitive situation better in terms of economic efficiency? (Hint: Compare the amounts of producer and consumer surplus in each.)

h) Suppose that the four firms recognize that they can cooperate and act as a joint monopoly. What will they do? Will this move increase or decrease economic efficiency?

CHAPTER 16

TRANSPORT COSTS AND BORDERS

After you read this chapter and work the problems at the end, you should be able to:

Define comparative advantage, tariff, quota, terms of trade, zero-sum game, positive-sum game, and negative-sum game.

Compute opportunity costs when given information about production possibilities.

Explain what the Hotelling model says about the behavior of political candidates.

Explain why traveling costs can create unexploited value.

This chapter looks at the effects of transport or traveling costs and political borders. The first two parts examine cases in which buyers have to expend time and effort to go to a seller, and how this cost affects behavior and economic efficiency. The third and fourth parts show that the existence of national borders is irrelevant in terms of economic efficiency. Such borders are not irrelevant in terms of political incentives, and as a result economists have often been critics of government trade policies.

A key argument of this chapter is that decreasing the costs of making transactions, whether it be by reducing transportation costs or lowering man-made barriers to trade such as tariffs, increases consumer choice and thus economic efficiency.

HOTELLING'S MODEL

Suppose that two owners of refreshment stands, George and Henry, are trying to decide where to locate along a stretch of beach. Suppose further that there are 100 customers located at even intervals along this beach, and that a customer will buy only from the closest vendor. Finally, assume that the beach is short enough so that total sales are independent of where the vendors locate.

Suppose that initially the vendors locate at points A and C in Figure 1. These locations would minimize the average traveling costs of the buyers and would result in each vendor getting one half of the business. However, this solution would not be an equilibrium. If George moved from point A to point B, he would keep all customers to his left, and get some of Henry's customers. For similar reasons, Henry would move toward the center, and in equilibrium, both vendors would locate together in the middle.

This story of the beach was first told a half century ago by Harold Hotelling, and is called Hotelling's model. Though it can give some insights into businesses decisions concerning location and product characteristics, the model has been more useful in explaining certain political phenomena. Instead of two refreshment stands along a beach trying to attract dollars from customers, consider two political candidates along the political spectrum trying to attract votes from voters. Only the candidate who attracts the most votes will win, and to attract votes, a candidate must locate nearer to more voters than his opponent. With these rules, there is a strong tendency for each candidate to move to the middle.

In American politics this tendency has a predictable consequence for presidential candidates, who must "sell" on two beaches. To gain the nomination, the candidate must

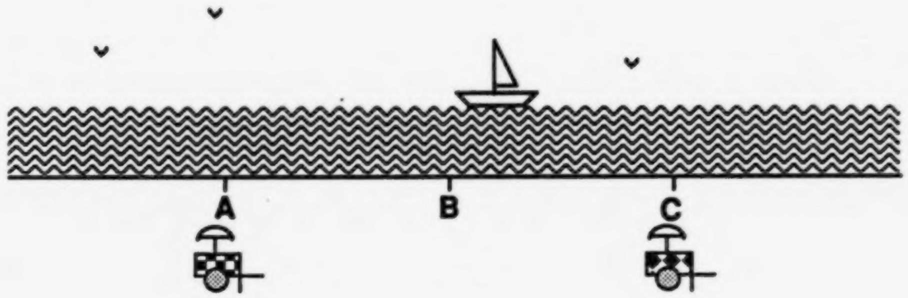


Figure 16.1 The Hotelling Beach.

position himself in the middle of the party. (However, see question 4.) Since the average Democrat has significantly different views than the average Republican, Republican and Democratic candidates sound quite different before nominations are decided. Once a candidate wins his party's nomination, he must "sell" to a different beach. Republican candidates move to the left and Democratic candidates move to the right. By election time their positions on issues usually sound close enough so that factors such as personality emerge as keys to the election.

There have been some notable exceptions to this pattern. In 1964 Barry Goldwater won the Republican nomination standing well to the right of the average voter, and was unable or unwilling to re-position himself in the center. In 1972 George McGovern won the Democratic nomination standing well to the left of the average voter, and was unable or unwilling to re-position himself. Both lost in landslides.

A problem with the Hotelling model when applied to commerce is that the results are very sensitive to the cost assumption. There must be some cost to traveling because customers prefer the closest vendor. But these costs must be small, because the people at the end of the beach continue to buy the same amount no matter how far they are from the nearest vendor. If traveling costs are less, then people might not care whether they go to the nearest vendor. If they are greater, so that when the vendor gets far away, people do not bother to go, the vendors will no longer cluster at the middle.

Suppose that the beach is a long beach, and people more than 1000 feet away from any seller buy nothing. Also assume that the beach is 4000 feet long, and the two vendors start at the middle. Originally George sells to customers located from the 1000 foot mark to the middle at 2000 feet, and Henry sells from 2000 feet to 3000 feet. If George moves to the 1000 foot mark, he will gain 1000 feet of new territory, and he will lose only 500 feet to Henry. At the 1000 foot mark, he will sell to all people from 0 to 1000 feet. He will also sell to those people between him and Henry who are closer to him. Since Henry did not move, but stayed at the 2000 foot mark, George will get all the customers up to the 1500 foot mark. Equilibrium in this case will occur only when Henry moves to the 3000 foot mark.

In Hotelling's original model with small traveling costs, location decisions were not economically efficient. By increasing traveling costs, it seems that we can have location decisions which are economically efficient. This is a rather strange result because the next section shows that, beginning with an efficient world of price takers, the addition of traveling costs results in economic inefficiency.

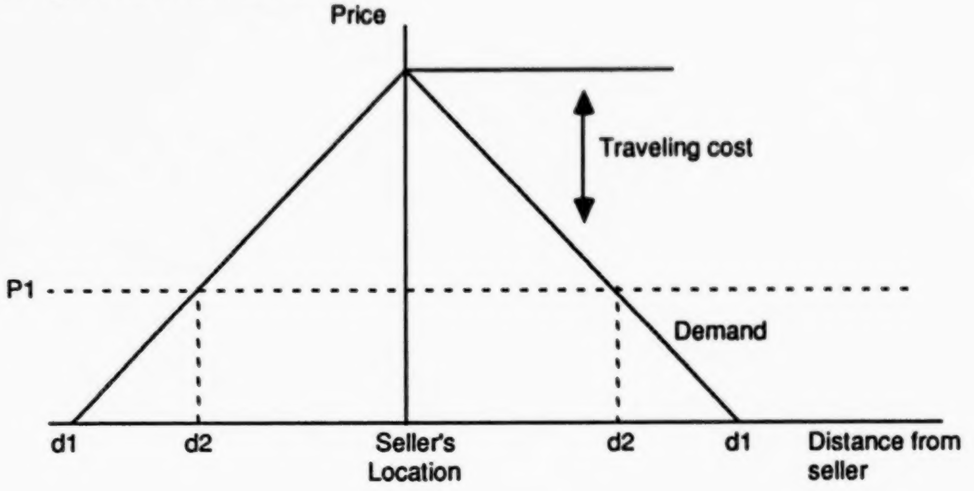
MONOPOLISTIC COMPETITION

In a world of many sellers, each of whom is a price taker, no seller can have a pricing strategy. If any seller tries to raise his price above the prevailing market price, he loses all sales. A world of price takers is, as Chapter 11 showed, economically efficient.

In the world around us there are few industries made up of price takers. An important reason for this fewness is that traveling to sellers costs buyers time and effort. As distance between customer and seller increases, these costs rise and the net benefit from buying from the seller drop.

Figure 2 illustrates these ideas by showing the demand "tent" which an isolated seller faces. Value is measured vertically and distance horizontally. In the center, where the seller is located, there are no traveling costs, so the value which buyers living there get from purchasing from this seller are as high as possible. As the distance from the center increases, so do traveling costs, and this causes the marginal benefit from buying from this seller to decrease. Beyond distance *d1* traveling costs exceed the benefit people receive from the product, and as a result people would not travel to the seller even if the product was free.

Figure 16.2 A demand tent.


Price
Traveling cost
P1
Demand
d1
d2
Seller's Location
d2
d1
Distance from seller

The seller, however, does not give away his product but charges for it. In deciding whether to buy, customers must compare the benefit they get from the product to both costs of traveling and the price they pay. The traveling costs are captured in the downward slope of the demand curves that make up the tent. In Figure 2 the seller charges price *p1*. Beyond distance *d2* the marginal benefit of purchasing at this seller, which is the value of the product less traveling costs, are less than price, and customers will not buy.

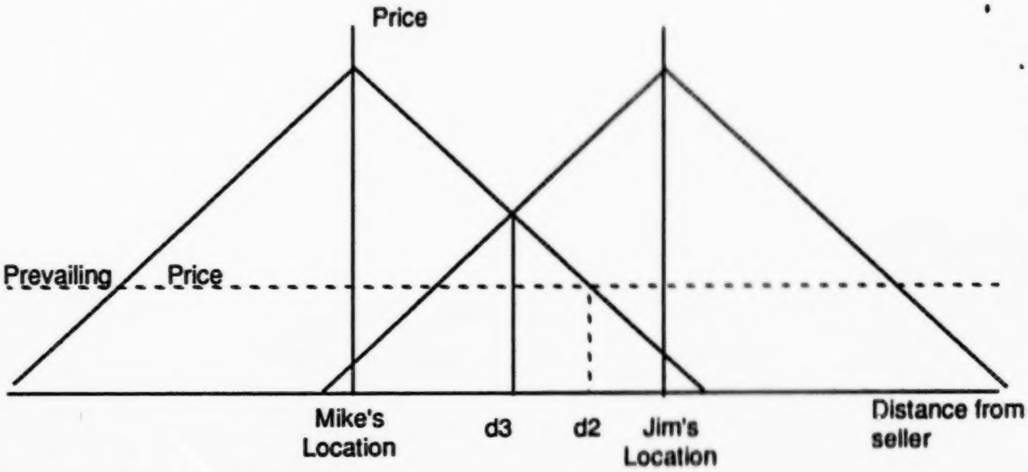

Price
Prevailing Price
Mike's Location
d3
d2
Jim's Location
Distance from seller

Figure 16.3 Overlapping tents.

Figure 16.4 Raising price cuts sales but does not eliminate them.

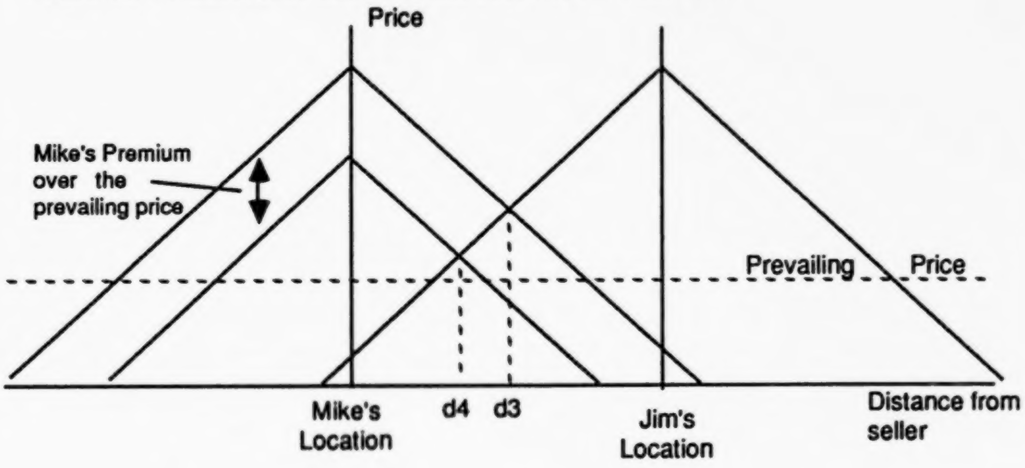


Figure 3 shows overlapping tents of two sellers, Mike and Jim. If both charge price *p2*, Mike will sell out to *d3*, not to *d2*. Some people located beyond *d3* would still receive net benefit from buying from Mike, but they can get more net benefit by buying from Jim. Recall that net benefit is the distance from the price line to the tent.

Figure 4 illustrates what happens if Mike raises his price by *d* above the prevailing prices of other sellers. The higher price is equivalent to a reduced benefit from buying from Mike, so Figure 4 shows his tent shifting down by distance *d*. Though Mike is charging more than other sellers, he does not lose all customers, but only sales from *d3* to *d4*. Because traveling costs give him some market power, he faces a downward sloping demand curve.

The model of monopolistic competition puts our situation into a more familiar form of demand and cost curves. Figure 5 shows a seller with a downward sloping demand curve and a conventional marginal cost curve. Because the demand curve slopes downward, the marginal revenue curve lies below it. The sellers maximizes profit by selecting that output at which marginal revenue equals marginal costs, and charges as much as he can, which is price *p2*. In the long run there can be no economic profit because there is free entry into the industry. If there are any profits, others will enter the industry, positioning themselves so as to take away customers from the most profitable sellers. The zero profit condition implies that at equilibrium average revenue (which is demand) must just equal average cost. When average revenue equals average cost, average profit is zero, and so total profit must also be zero.

The model of monopolistic competition was considered important when it was introduced for two reasons. First, the situation it described seemed the most common form of industry. Both the single sellers of monopoly and the many sellers of price-taking competition are uncommon in comparison. Furthermore, monopolistic competition describes more than traveling costs in a geographical sense. The distance between sellers can be in the minds of buyers. Product differentiation, which results in many products which are similar but not identical, also creates a distance between products, and it was this distance that seemed especially important to the developers of monopolistic competition.

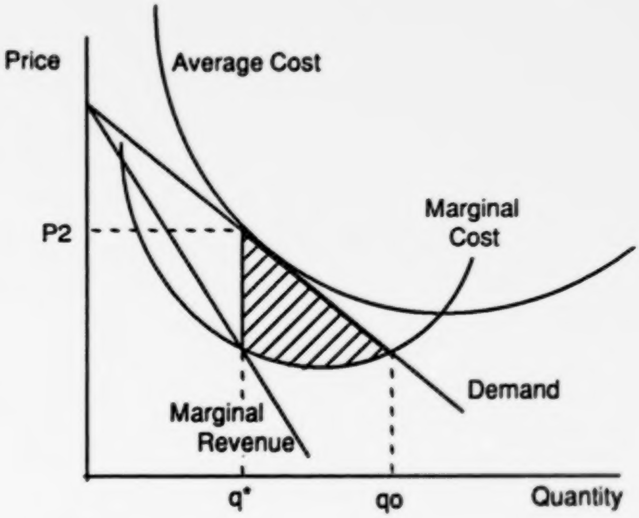


Figure 16.5 Monopolistic Competition.

(Notice that while the theory of monopolistic competition suggests that there may be too much product differentiation, Hotelling's model suggests that there may not be enough.)

Second, notice that, because price exceeds marginal cost, Figure 5 contains an area of welfare loss. There is unexploited value which neither firms nor customers obtain. Because monopolistic competition was both common and economically inefficient, one could argue that market systems were inherently inefficient.

However, the welfare loss in the case of monopolistic competition may be illusionary. Firms could obtain this value if they could price discriminate, selling beyond *q** up to *qo*. If they do not do so, then the resources needed to obtain the information required to price in this fashion must be more valuable than the triangle of unexploited value. Or consider possible government solutions. The problem is that sellers are too small and charge too high a price. The government could react by limiting the number of sellers and forcing them to charge lower prices. But this policy would increase traveling costs of buyers. The market may not give the optimal number of sellers (and hence the optimal distance between them), but the cost of gathering the information that would let government decide the matter is almost certainly greater than any possible welfare gain (even ignoring the political incentives which always go with government solutions).

Some economists have argued that the distance between products is often phoney, that firms differentiate products to fool consumers. Their argument is surely correct in some cases. In other cases, however, product differentiation probably exists because it reflects differences in people's tastes. Again, there is a cost to deciding whether in each case product differentiation manipulates buyers or caters to them. It is not clear that any government policy which tries to eliminate "bad" differentiation will have benefits which exceed its costs.

When the cost of correcting a problem exceeds the possible gain from the correction, is there any real welfare loss by allowing the problem to continue? The model of monopolistic competition shows that real market systems fall short of theoretical constructs which assume away the problems of information and of making agreements among people. But any real-world system looks bad compared to theory which assumes away problems.

Finally, the theory developed above does have at least one interesting use in explaining the real world. If traveling costs are reduced, the tent of each seller gets less steep, which means that the demand curve facing each seller gets more elastic. As a result, the marginal revenue curve shifts upward, and will intersect the marginal cost curve at a higher quantity (or greater distance). For individual firms to expand sales when the industry sales are constant requires that some firms disappear. Reducing traveling costs reduces the number of firms. In the United States development of the automobile and the highway system drastically cut traveling costs. As a result, throughout America there are thousands of villages which have completely or largely lost their retail districts.

INTERNATIONAL TRADE

Most economists from the time of Adam Smith have favored free international trade and have opposed tariffs and quotas. A tariff is a tax on an import and a quota is a limitation on the number of units that may be imported. On the other hand, a great many people over the years have argued for some form of restriction on international trade, for "protection" from foreign producers. Politicians have at times been persuaded by economists and others who have been in favor of free trade, and at other times by those arguing for restrictions. The United States and other Western nations have had sizable fluctuations in the average level of tariffs.

Those who favor protection sometimes claim that it is unfair to ask Americans to compete with cheap foreign labor. Some go on to argue that because labor is cheaper abroad, foreigners can undersell the United States in any product, and that if American business is not protected, the very standard of living in the United States is in danger. Economists have written some simple yet effective stories to rebut this argument. One of the best and oldest is Frederic Bastiat's "Petition of the Candlemakers."

Bastiat's petition, published in 1874 in English, is a satire arguing that France would be better off if it had a law forbidding buildings to be lit with natural light. Without windows, Bastiat tells the reader, there would be a huge demand for artificial light, and a great many people would benefit from this demand. All the processes of candlemaking, including agriculture because tallow comes from sheep and cattle, would benefit. The maritime industry would benefit because more ships would put to sea to hunt whales. Metal workers would be needed to make light fixtures. Bastiat argues that directly or indirectly all Frenchmen would be better off because of the greater production of artificial light.

Bastiat takes a position which everyone should recognize as absurd and shows that it makes sense in terms of the arguments used to support tariffs. He carefully ignores the concept of opportunity cost, but those who support tariffs never mention this concept either. If a country wants to reject cheap light from the sun, or cheap food from foreign nations, it will need resources to satisfy the demand for these items. Using resources to produce artificial light or domestically grown food means that these resources cannot be used for other purposes.

A fable by James Ingram also illustrates that free international trade is production efficient. In his book, *International Economic Problems* (John Wiley, 1970), he tells of a mysterious entrepreneur, Mr. X., who announces to the world that he has found several amazing discoveries which allow him to produce cheap televisions, automobiles, cameras, and other goods. He sets up a plant on a large tract of land along the coast of North Carolina, hires 5000 employees who

TABLE 1
OPPORTUNITY COST OF FISH AND COCONUTS

FOR FRIDAY:		
	1 FISH COSTS 1 COCONUT	1 COCONUT COSTS 1 FISH
FOR CRUSOE:		
	1 FISH COSTS 2 COCONUTS	1 COCONUT COSTS 1/2 FISH

are sworn to secrecy, and begins buying grain, coal, and machinery. As the trains of grain and coal roll into his factory, other trains full of televisions and automobiles roll out of his factory into showrooms across the country. Mr. X is hailed as another Edison or Bell, and his company becomes a favorite with Wall Street investors.

Consumers love Mr. X because his products are so much cheaper than what they could buy before. Of course his competitors dislike him, but their attempts to get laws restricting his operations get nowhere. The Houses of Congress ring with speeches saying that some economic adjustment is an inevitable by-product of technological progress.

Then one day a small boy trying out his new skin-diving gear accidentally penetrates Mr. X's security shield and learns Mr. X's secret. Nothing is produced at the factory. It is all a front for a giant import-export business. Mr. X transforms grain and coal into autos and televisions by trade. His secret revealed, Mr. X is reviled and his factory shut down. Congressmen proclaim that the American standard of living has had a narrow escape from the threat of cheap foreign labor and urge more money for research in industrial technology.

The reasons that free trade is desirable can be developed by extending the discussion of the Crusoe economy in Chapter 2. Suppose that Robinson Crusoe, living on an almost deserted desert isle, can either catch four fish a day or find eight coconuts. One day he discovers that Friday also lives on the island. If Friday can either catch six fish a day or find seven coconuts, can Crusoe and Friday profit by specialization and trade? The answer is clearly "yes." Crusoe is the better coconut gatherer, and Friday the better fisherman.

However, suppose that Friday can either catch ten fish or find ten coconuts. Friday is now better than Crusoe in both activities. Can there be mutual benefit from trade in this case? Or should Friday do all the work and Crusoe none? Or should Friday refuse to trade since he is better in both? It was a major achievement of David Ricardo early in the 19th century to show that in this second Crusoe-Friday story both parties could benefit from trade. His results contributed to the long reign of relatively free trade in 19th century England, and thus to the prosperity that England enjoyed in this period.

To see that mutually beneficial trade is possible even though Friday is better, in all activities, one must look to opportunity costs. Individually, both Friday and Crusoe trade with nature in the production process. Crusoe can get another fish only by giving up time in which he could find two coconuts, and in getting another coconut he sacrifices one half a fish. Thus a fish costs Crusoe two coconuts and a coconut costs one half a fish. Friday can get another fish by giving up the time during which he can find another coconut, or one tenth of the day. During this time, he could find one coconut. For Friday trading with nature means that one fish costs one coconut and vise versa. Looking at these opportunity costs tells us that Crusoe finds coconuts cheaper and Friday finds fish cheaper. Table 1 summarizes these results.

TABLE 2
PRODUCTION POSSIBILITIES

CRUSOE		FRIDAY	
fish	coconuts	fish	coconuts
4	0	10	0
3	2	9	1
2	4	5	5
1	6	3	7
0	8	0	10

We still have not discovered whether Friday and Crusoe could do better trading with each other rather than with nature. A way to answer this question is to try a few prices. Suppose that one fish was worth one half a coconut, or one coconut was worth two fish. With this trading ratio both would find fish cheap and coconuts expensive. Therefore, both would want to sell coconuts and buy fish. Hence at this price no trading would take place.

Suppose the trading ratio was one fish for 1.8 coconuts (and thus one coconut cost 5/9 fish). With this ratio Crusoe would find fish cheap—rather than spend two coconuts to catch one, he could spend 1.8 coconuts and buy one. Hence Crusoe would be willing to sell coconuts. Friday would find coconuts cheap—rather than give up one fish by gathering his own coconuts, he could sell 5/9th of a fish and get one. Hence Friday would be willing to sell fish. Trade will take place because both individuals find that it improves their well-being.

In the above example, trade occurs because of **comparative advantage**. Friday is better in everything than Crusoe, but he is "more better" in catching fish and "less better" in finding coconuts. Crusoe is worse than Friday in everything, but he is "less worse" in finding coconuts. Though they both benefit from trade, Friday will maintain a higher standard of living.

The example of Crusoe and Friday also illustrates that exchange is not a zero-sum game, but a **positive-sum** game. In a zero-sum game whatever anyone wins comes at someone else's expense. In poker, for example, if one person wins $100, some other person(s) must have lost $100. If the amount of winnings is added up and the amount of losings is subtracted away, the result will be zero. The term "zero-sum game" reflects this total. In contrast, both parties in a voluntary exchange can benefit. Because total winnings exceeds any losses, the name "positive-sum" game is appropriate. A **negative-sum** game is also possible. Here the winnings will be less than losses. War is an example of a negative-sum game.

A reason that few people prior to Adam Smith seem to have recognized this mutually beneficial aspect of exchange may be that in his day much exchange involved bargaining. In bargaining the seller tries to get as high a price as he can and the buyer as low a price as possible. If the seller can get $2.50 for a product rather than $2.00, he benefits from the higher price and the buyer is worse off than he would be at the lower price. People do discuss who got the "better of the bargain." This feature of exchange, of vital importance to those involved in a market, can obscure the fact that no exchange will take place unless both parties believe that they benefit from it. Bargaining determines how big the producer's surplus will be relative to the size of the consumer's surplus, but unless both buyer and seller each have some surplus, no trade will take place.

Crusoe and Friday could be replaced by two nations. The principle of comparative advantage continues to hold, and it implies that the world as a whole will not be operating on its production-possibilities curve—that it will be production inefficient—if each nation is self-sufficient.

This inefficiency can be illustrated with the numbers given in Crusoe-Friday story to produce the production-possibilities tables in Table 1. Suppose in self sufficiency Crusoe chooses two fish and four coconuts and Friday chooses five fish and five coconuts. Then total island production is seven fish and nine coconuts. But with total specialization, with Friday producing only coconuts and Crusoe producing only fish, island production would be eight fish and ten coconuts, which means that under self sufficiency island resources were used inefficiently. Or suppose that originally Crusoe was self sufficient at one fish and six coconuts and Friday at five fish and five coconuts. Island production is six fish and eleven coconuts. In this case only partial specialization might be desirable. If Friday produces three coconuts and Crusoe eight, island production could be eleven coconuts and seven fish, which is a gain of one fish compared to production with no specialization.

BARRIERS TO TRADE

If the gains from free trade are clear theoretically, why are there tariffs and quotas? One important reason is that the costs of tariffs are defused and hard to see, while the benefits are concentrated and easy to see. A tariff or quota raises the price of a product. This will harm consumers of the product, but will benefit the domestic producers of the product. If the harm done to consumers is spread over a great many, and if the exact amount of harm done to each is unclear and small, they may do little lobbying to protect themselves. (Remember the rationally ignorant voter.) If the benefits to producers are clear and concentrated on relatively few, the producers will lobby for protection. Thus the pressures that political leaders face may not reflect the ratio of total costs and benefits of the action.

However, the fact that there are both winners and losers to any change in tariffs and quotas raises a problem. The nature of this problem can be illustrated with a numerical example. Table 3 gives three distributions of four fish and seven coconuts between Crusoe and Friday, and the value Crusoe and Friday put on another coconut or fish. The condition of exchange efficiency from Chapter 11 indicates that distribution *A* is not efficient. Friday places much more value on another fish than he places on coconuts, while Crusoe considers them of equal value. Both Friday and Crusoe can gain from trade.

Distributions *B* and *C* are both exchange efficient. In both distributions Crusoe and Friday consider fish twice as valuable as coconuts. No rearranging of fish and coconuts can take place without harming one.

Moving from distribution *A* to distribution *B* is a move from an inefficient distribution to an efficient one which might happen with voluntary exchange. In distribution *A* Crusoe thinks fish and coconuts have the same value, so he should be willing to give up .7 fish to get another coconut. In distribution *A* Friday thinks fish are five times as valuable as coconuts, so he should be willing to give up one coconut to get .7 fish. After the trade, the law of diminishing marginal utility would change valuations. Crusoe has less fish, so he values another unit of fish more, and more coconuts, so he values another unit of coconuts less.

A move from distribution *A* to distribution *C* is also a move from and inefficient to an efficient distribution. However, this move would not take place voluntarily. Crusoe's well-being is improved, but Friday's has decreased. This is the sort of change which occurs when a tariff is removed. (Imposing a tariff involves a move similar to a move from distribution *C* to distribution *A*.) A person claiming that distribution *C* is better than distribution *A* implies that the gain to Crusoe is of greater value or importance than the

TABLE 3
THREE DISTRIBUTIONS OF GOODS

Distribution		*Situation*	*Value of fish in terms of coconuts*
A	Crusoe:	3 fish 3 coconuts	1 fish = 1 coconut
	Friday	1 fish 4 coconut	1 fish = 5 coconuts
B	Crusoe	2.3 fish 4 coconuts	1 fish = 2 coconuts
	Friday	1.7 fish 3 coconuts	1 fish = 2 coconuts
C	Crusoe	3.3 fish 5 coconuts	1 fish = 2 coconuts
	Friday	.7 fish 2 coconuts	1 fish = 2 coconuts

loss to Friday. This seems arbitrary. To make it seem less arbitrary, some economists have suggested that if the gain to Crusoe is large enough so that he could pay Friday enough to reimburse him for his loss, then the change was desirable even if payment is not made.

This problem of change harming some and helping others is not confined to the issue of tariffs and quotas. Any non-trivial change involving an outward shift in the production-possibilities curve in a complex social system will aid some and harm others. The development of the automobile harmed those who produced buggies. The development of the electronic calculator harmed those who produced slide rules and mechanical calculators. In England between 1811 and 1820, organized bands of craftsmen called "Luddites" protested the threat of new techniques by smashing textile machinery which was displacing them. The list of examples of changes harming some and helping others is very long. The destructive nature of change was captured in the phrase that Joseph Schumpeter used to describe the most noteworthy aspect of capitalism: "the process of creative destruction." Because so many changes involved both groups that are harmed and those who are helped, unless one is willing to make some interpersonal comparisons, the concept of economic efficiency is of little use in making policy judgements.

One could argue for tariffs and restraints on trade by acknowledging that a few will be helped and many harmed, and arguing that the well-being of the few should be given more importance than the well-being of the many. But no one ever publicly argues for protection in this way. Most public arguments for protection are simply wrong because they do not acknowledge that anyone will be hurt. However, there are a few special cases in which an appealing and logically correct argument for protection can be made.

One such argument is the "infant industry" argument. A country may have a comparative advantage in a product once the industry is established, but it may need protection to build up to the point at which it can successfully compete. Though sound on paper, this argument is subject to considerable abuse in practice. It is difficult to forecast which industries will actually be able to compete successfully when mature, and hence many unneeded tariffs may be enacted. Further, once a tariff is in place, it may be difficult to repeal even when it is no longer needed for "infant" protection. No one gives up government favors without a fight.

Two other arguments rely on a division of the world into us-them categories, a division which political borders encourage. People do make these divisions—most people find that news about disasters affecting one's countrymen are more distressing than the same disasters affecting foreigners, even though both are complete strangers. Though economics rarely incorporates these divisions into its analysis, politicians respond to them.

The first of the us-them arguments is a national defense argument. If foreign competition threatens a skill or industry which would be necessary during wartime, the government may decide that protecting the skill or industry is necessary. The second argument notes that though the total harm to transactors will exceed the amount of revenue that a tariff brings to the government as tax revenue, the harm is borne both by domestic buyers and foreign sellers. Under certain conditions the foreign sellers may bear enough of the harm so that the loss to the domestic buyers is less than the amount that the government collects. In this case it may be in the country's (though not in the world's) interest to establish a tariff. However, this type of situation may be relatively uncommon in the real world since it requires that the nation imposing the restrictions buy or sell a great proportion of the item that is restricted.

SUMMARY

This chapter examined exchange when distance and political borders exist. It looked at the explanations of the effects of distance which the Hotelling model and the model of monopolistic competition offer. In examining the effect of political borders, the chapter showed that restrictions on trade lead to production inefficiency, and also cause a redistribution of income.

QUESTIONS

1. Suppose there are three sellers on the Hotelling beach. If initially each has one third of the customers, where will they be located? Can anyone increase his share by moving? Why is the middle a poor place to begin? If the middle wants to survive, what must he eventually do? Can you see any equilibrium if all three continue to exist?

2. Suppose the Dukedom of Scimonoce is divided into two separate parts and no trade is permitted between them. West Scimonoce is made up of plots A and B and East Scimonoce is made up of plots C, D, and E. (See Question 9 of Chapter 2.) If people in each section want 10 units of oats produced, how much corn will be produced in West Scimonoce? In East Scimonoce? What is total oat and corn production? Where is this on the graph in question 9 of chapter 2?

3. Though Adam Smith advocated free trade, David Ricardo was the first to clearly explain the idea of comparative advantage. For his explanation he began with production costs similar to those below.

ENGLAND	PORTUGAL
1 barrel of wine requires	1 barrel of wine requires
1 man-day	2 man-days
1 bolt of cloth requires	1 bolt of cloth requires
1 man-day	3 man-days

If it took three man-days to produce a barrel of wine in England instead of one, there would clearly be gains from trade. England would then have a **absolute** advantage in cloth, and Portugal an absolute advantage in wine. Ricardo, however, said that with the numbers above, even though England could produce both items with less labor, there were gains from trade. Let's see if you can recreate his reasoning below.

a) Using these data, compute the opportunity costs of wine and cloth below:

In England:	In Portugal:
to get 6 barrels of wine, it must give up ___ bolts of cloth.	to get 6 barrels of wine, it must give up ___ bolts of cloth.
to get 6 bolts of cloth, it must give up ___ barrels of wine.	to get 6 bolts of cloth, it must give up ___ barrels of wine.

(Hint: Producing 6 barrels of wine requires how many man-days? If these man-days are taken away from cloth production, how much cloth production is lost? (Using 6 keeps fractions out of the results and makes computation easier.))

b) Based on the opportunity costs you computed, which country produces wine cheaper? Which cloth?

c) Suppose 6 barrels of wine trade for 5 bolts of cloth. (This is the same as 6 bolts of cloth trading for 7 1/5 barrels of wine.) Would England find it cheaper to buy wine or produce it? To buy cloth or produce it? Would Portugal find it cheaper to buy wine or produce it? To buy cloth or produce it? Could mutually advantageous trade take place?

d) Suppose that 6 barrels trade for 7 bolts. Would England find it cheaper to buy wine or produce it? To buy cloth or produce it? Would Portugal find it cheaper to buy wine or produce it? To buy cloth or produce it? Could mutually advantageous trade take place?

4. Portlandia produces each year either sails and/or boats as follows:

	A	B	C	D	E
SAILS	0	20	40	80	100
BOATS	75	60	45	15	0

Starboardia produces each year either sails and/or boats as follows:

	A	B	C	D	E
SAILS	0	8	16	24	48
BOATS	36	30	24	18	0

a) Prior to trade, Portlandia is at D (that is, it produces 80 sails *and* 15 boats) and Starboardia is at C. Note that both have straight-line production-possibilities curves, which means that opportunity cost is constant. Are comparative costs such that trade will occur?

b) If so, which country will produce sails?

c) What are the gains from specialization?

d) What are the limits to the terms of trade?

e) What is the opportunity cost of a boat in Portlandia? In Starboardia?

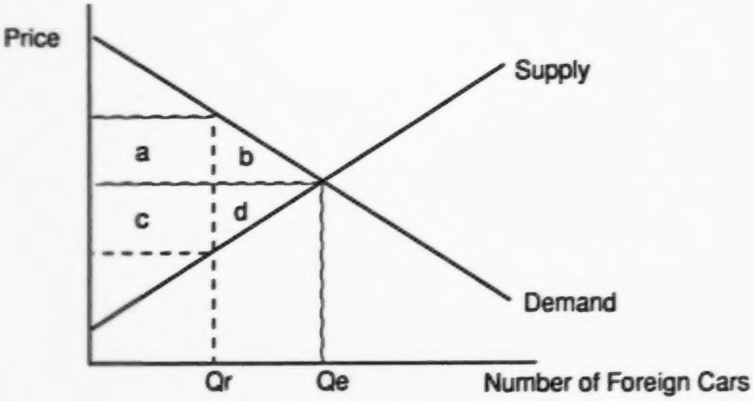
f) Which of the following firms could exist in Portlandia? A firm that can always make either:

a. 2 sails or 3 boats a year.
b. 2 sails or 1 1/2 boats a year.
c. 3 sails or 4 boats a year.
d. 5 sails or 6 boats a year.
e. all of the above
f. none of the above.

5. Dawn Key is a very successful businesswoman. She is also skilled in the domestic arts. Though she could clean her house in two hours, she prefers to hire a cleaning lady who spends five hours doing the same job. Though she is a skilled cook, when she hosts a party, she hires a caterer. The caterer takes ten hours to prepare the food, while Ms. Key could do the same work in only six. Why do Ms. Key's actions make sense in terms of comparative advantage?

6. The price ratio of fish to coconuts used in discussing trade between Crusoe and Friday is called the **terms of trade**. There is a range of price ratios that will work so that trade will take place. What are the limits of this range in the example used in Table 1?

7. In order to help domestic automobile producers during the 1970s, the U.S. government sought ways to restrict sales of foreign automobiles, especially those made in Japan. In terms of a supply-demand diagram, the market would have allowed in *Qe* automobiles, but the government wanted only *Qr* to be sold. A restriction that accomplishes this goal will redistribute parts of the consumer and producer surpluses.



(The parts that will be affected are labeled *a*, *b*, *c*, and *d* on the graph.)

a. One way to restrict is with a tariff. What would happen to areas *a*, *b*, *c*, and *d* with a tariff?

b. Another way to restrict is with a quota. A quota requires that the right to import be rationed in some way. Suppose that the government gave the right to import to domestic automobile companies. How would such a program have affected the four areas in the graph?

c. The government neither established a tariff nor a quota. Rather it negotiated an agreement with the Japanese by which the Japanese automobile manufacturers voluntarily limited their sales to the U.S. How would such an agreement affect the four areas?

d. While Ronald Reagan was president, the agreement with the Japanese expired and he did not renew it. The Japanese automobile manufacturers, however, decided to continue limiting their sales to the U.S. Does this action indicate that the Japanese are poor businessmen?